Title: Data-driven model of glycolysis identifies the role of allostery in maintaining ATP homeostasis

Authors: Mangyu Choe[1,2], Tal Einav[4,6], Rob Phillips[4,5], Denis V. Titov[1,2,3]*.

Affiliations:

[1]Department of Nutritional Sciences and Toxicology, University of California; Berkeley, CA 94720, USA.

[2]Department of Molecular and Cell Biology, University of California; Berkeley, CA 94720, USA.

[3]Center for Computational Biology, University of California; Berkeley, CA 94720, USA.

[4]Division of Biology and Biological Engineering, California Institute of Technology; Pasadena, CA 91125, USA.

[5]Department of Physics, California Institute of Technology; Pasadena, CA 91125, USA.

[6]Basic Sciences Division and Computational Biology Program, Fred Hutchinson Cancer Research Center, Seattle, WA 98109, USA.

*Correspondence: titov@berkeley.edu

SUMMARY

Glycolysis is a conserved metabolic pathway that produces ATP and biosynthetic precursors. Here, we use a mathematical model to investigate how the control of glycolytic enzymes through allostery and mass action accomplishes various tasks of ATP homeostasis, such as controlling the rate of ATP production, maintaining high and stable ATP levels, and ensuring that ATP hydrolysis generates energy. Our model uses enzyme rate equations with >150 biochemical parameters estimated from experiments to accurately predict glycolysis activity in living cells. We find that allostery is required for the robust maintenance of high and stable ATP levels by preventing a futile cycle between ATP-producing and ATP-consuming enzymes of glycolysis, while mass action alone is sufficient for regulating the rate of ATP production and ensuring that ATP hydrolysis generates energy. Our results suggest that the functions of allostery are distinct from its postulated role in controlling metabolic pathway rates.

Keywords: glycolysis, metabolism, ATP homeostasis, allosteric regulation, Monod-Wyman-Changeux model, robustness, systems biology, kinetic modeling.

INTRODUCTION

Glycolysis is conserved across all domains of life. This key pathway harnesses the breakdown of glucose to produce energy in the form of ATP and precursors for the biosynthesis of amino acids, nucleotides, carbohydrates, and lipids. Glycolysis produces ATP both directly by a process referred to as fermentation or aerobic glycolysis (Figure 1A) and indirectly by producing pyruvate that is used as a substrate for ATP generation by respiration. The net reaction of aerobic glycolysis converts extracellular glucose to extracellular lactic acid, while the only net intracellular reaction is the production of ATP from ADP and inorganic phosphate (Figure 1A). Aerobic glycolysis is the most active metabolic pathway in proliferating mammalian cells (an observation known as the Warburg Effect[1]), where it can satisfy all the ATP demand in the absence of respiration[2,3]. The reliance of proliferating cells on aerobic glycolysis for ATP production and the lack of intracellular products except for ATP makes glycolysis a convenient self-contained system for studying ATP homeostasis.

The regulation of ATP homeostasis by glycolysis is incompletely understood. The rate of ATP production by glycolysis is postulated to be controlled by allosteric regulation of hexokinase (HK), phosphofructokinase (PFK), and pyruvate kinase (PK) by ATP, ADP, AMP, inorganic phosphate, and other regulators. However, the latter mechanism has never been validated. In fact, most of the allosteric regulators of HK, PFK, and PK are also substrates and products of glycolytic enzymes, making it uncertain if their effect on the rate of glycolysis is driven by allostery or mass action. Furthermore, in addition to regulating the rate of ATP production, glycolytic ATP homeostasis performs other tasks, such as ensuring that the ATP hydrolysis reaction generates energy and maintaining most of the adenine nucleotide pool in the form of ATP. The regulatory mechanisms that allow glycolysis to perform the latter tasks are also unknown.

Due to the complexity of glycolytic ATP homeostasis, a better understanding of its regulation requires the use of mathematical modeling. This has been an active area of research for several decades, with many groups developing mathematical models of glycolysis[4–17]. These efforts range from simple idealized models involving a subset of enzymes[4–6,8,12] to full-scale models containing all the enzymes[7,9–11,13–17]. Previous studies had three key limitations that we addressed in this report. First, the regulation of tasks of glycolytic ATP homeostasis other than the rate of ATP production has not been investigated. Second, a systematic analysis of the role of mass action and allosteric regulation in controlling glycolysis has not been performed. Third, enzyme rate equations used in the published models were not systematically derived using comprehensive data from multiple *in vitro* kinetic studies.

Here, we investigated the regulation of glycolytic ATP homeostasis using a combination of theory and experiments. The core of our approach is a mathematical model that uses enzyme rate equations to describe glycolysis activity using a system of ordinary differential equations (ODEs). The key components of our model are the newly derived rate equations for allosterically regulated enzymes HK, PFK, glyceraldehyde-3-phosphate dehydrogenase (GAPDH), and PK. We used the Monod-Wyman-Changeux model[18–22] in combination with rigorous statistical approaches to identify the relevant rate equations and estimate the kinetic constants using a manually curated dataset of thousands of data points from dozens of publications. We solved a difficult problem of parameter estimation of a large mechanistic model by fully constraining our model with 157 parameters estimated independently based on proteomics, metabolomics, and *in vitro* kinetics data from hundreds of datasets. Our model robustly performed the key tasks of glycolytic ATP homeostasis and accurately predicted absolute concentrations of glycolytic

intermediates and isotope tracing patterns in live cells. Analysis of our model showed that allosteric regulation is required for the maintenance of high ATP levels such that most of the adenine nucleotide pool is maintained in the form of ATP. At the same time, mass action performed other tasks of glycolytic ATP homeostasis previously attributed to allostery, such as controlling the rate of ATP production to match ATP supply and demand, and maintaining ATP, ADP, and inorganic phosphate levels such that ATP hydrolysis reaction generates energy. Our results suggest a revision of the role of allostery in regulating glycolysis and likely other metabolic pathways.

RESULTS

Tasks of glycolytic ATP homeostasis

Glycolytic ATP homeostasis performs multiple tasks that likely require distinct regulation. In this section, we highlight three tasks of glycolytic ATP homeostasis that we have investigated (Figure 1A):

Task #1: matching the rates of ATP production and consumption. Any mismatch between the glycolytic ATP production rate and the ATP consumption rate by intracellular processes can only be short-lived due to the ~1-100 seconds turnover time of intracellular ATP[23]. The rate of glycolytic ATP production is postulated to be controlled by allosteric regulation of HK, PFK, and PK.

Task #2: ensuring that ATP hydrolysis generates energy. The energy of ATP hydrolysis is set by the extent to which ATP hydrolysis reaction is maintained far from equilibrium, which, in turn, is determined by the relative levels of ATP, ADP, and inorganic phosphate. Inside the cell, ATP hydrolysis reaction is maintained 10^9-10^{11}-fold away from equilibrium or 20-25 k_BT of energy[23]. It is not well understood how glycolysis controls ATP, ADP, and inorganic phosphate levels to keep ATP hydrolysis out of equilibrium.

Task #3: maintaining most of the adenine nucleotide pool in the form of ATP. Maintenance of the adenine nucleotide pool in the form of ATP under different ATP consumption rates is a well-established task of ATP homeostasis that ensures that ATP levels are high and stable[24,25]. This task is distinct from Task #2 as the ATP hydrolysis reaction can generate the same amount of energy at a wide range of ATP/ADP ratios due to the offsetting effect of inorganic phosphate. One function of maintaining high and stable ATP levels could be to ensure that ATP-consuming enzymes are not kinetically limited for ATP. For example, the K_M values for ATP of kinases, adenylate cyclases, myosin, Na^+/K^+-ATPase, Ca^{2+}-ATPase are in the 1-100 μM range–well below the 1-10 mM range of intracellular ATP levels. It is not well understood how glycolysis ensures that most of the adenine nucleotide pool is maintained in the form of ATP.

In addition, glycolysis performs the tasks above at a wide range of ATP demands, responds quickly to changes in ATP demand, and is robust to the physiological variability of enzyme levels, cofactor pool sizes, extracellular glucose and lactate concentrations, as well as kinetic constants that can change with temperature, pH, or growth conditions. Such robustness is an important property of the homeostasis[26].

Kinetic model of glycolytic ATP homeostasis

To investigate the regulation of ATP homeostasis, we have developed a comprehensive mathematical model of glycolysis. In this section, we provide an overview of the model development while most of the technical details are described in STAR Methods as well as comprehensive Notes S1-S3 that contain all the enzyme rate derivations. Our model includes key allosteric regulators based on a thorough literature search. Most glycolytic enzymes are encoded by several homologous genes that produce enzyme isoforms with tissue-specific expression and distinct kinetic and allosteric properties. To facilitate experimental testing, we chose to focus on glycolytic enzyme isoforms that are most abundant in proliferating cell lines (i.e., HK1, PFKP, and PKM2 isoforms of HK, PFK, and PK). We sought to develop a core model of glycolysis that supports ATP homeostasis in the absence of input from other pathways, and hence we only included allosteric regulators that are products or substrates of glycolytic enzymes. Our model converts a qualitative schematic of allosteric regulation (Figure 1B) into a precise mathematical language. As input, the model uses *i*) the extracellular concentrations of glucose and lactate, *ii*) cellular concentrations of cofactor pools and glycolytic enzymes, and *iii*) the rate of cellular ATP consumption (Figure 1C). We highlight these inputs separately from the model as these inputs are not controlled by glycolysis but depend on the cell type or experimental conditions and thus cannot be predicted by a glycolysis model. The model takes these inputs and uses enzyme rate equations assembled into a system of coupled ODEs to calculate the concentration of every glycolytic intermediate and the rate of every reaction. Our model can calculate both steady-state behavior and dynamical responses to perturbations. In effect, our model uses *in vitro* enzyme kinetics to predict any measurable property of glycolysis and many properties that cannot currently be measured.

In total, our model contains 157 parameters, including kinetic and thermodynamic constants and estimates of enzyme and cofactor pool concentrations, yet all these parameters are tightly constrained by experimental data. Kinetic constants were estimated from *in vitro* enzyme kinetics data as described briefly below and in full detail in Notes S1-S3. Thermodynamic constants are taken from the eQuilibrator database[27]. Enzyme and cofactor pool concentrations are estimated from proteomics and metabolomics data, respectively, as described in STAR Methods. We numerically simulate the model using DifferentialEquations.jl library[28] written in the Julia Programming Language[29]. Our code is heavily optimized so that it takes ~1-10 milliseconds to calculate the results of the model under given conditions using a single core of a modern computer processor. Optimized code allows us to run the model under millions of different conditions to systematically explore the regulation of glycolytic ATP homeostasis.

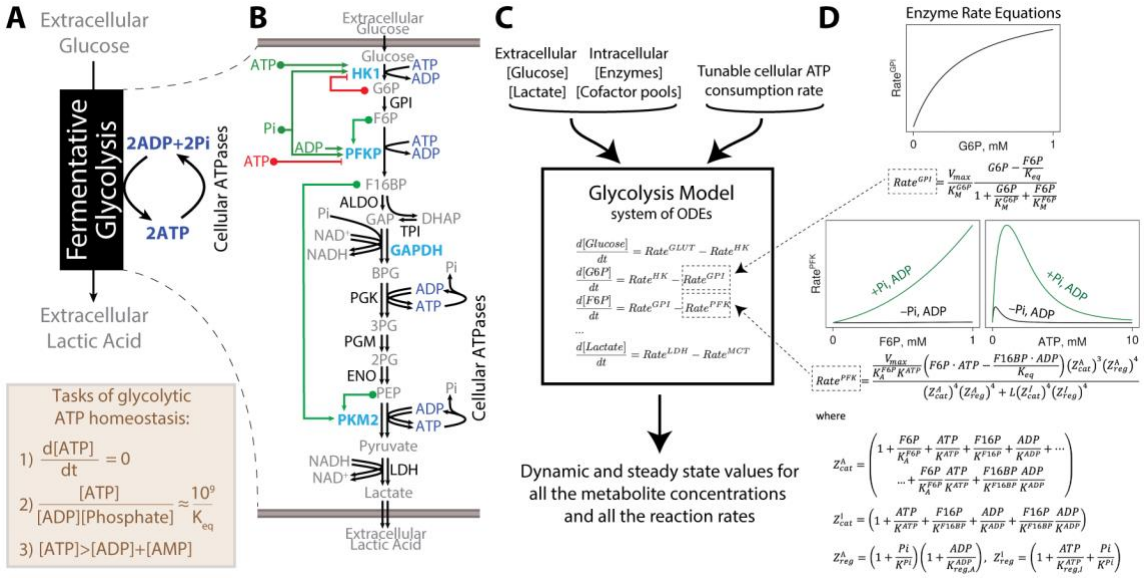

Figure 1. Overview of the mammalian glycolysis model. (A) Coarse-grained description of aerobic glycolysis highlighting its main function to transform glucose into ATP and listing the tasks of glycolytic ATP homeostasis. **(B)** Qualitative schematic of glycolysis showing the chain of enzymes (allosterically regulated enzymes in teal and the other enzymes in black) that convert substrates into products (gray). Allosteric activators (green) and inhibitors (red) that are included in the model are highlighted. **(C)** Schematic of the glycolysis model, including inputs and outputs. **(D)** Kinetic rate equations are shown for GPI and PFK, and plots are actual GPI and PFK rates calculated by the respective equations. Note the dramatic allosteric activation of the PFK rate in the presence of inorganic phosphate (Pi) and ADP. See also Notes S1-S3 for details on enzyme rate equation derivation.

The defining feature of our model is the use of newly derived kinetic rate equations to describe the activity of four allosterically regulated enzymes (i.e., HK1, PFKP, GAPDH, PKM2) using the Monod-Wyman-Changeux (MWC) model. MWC is a powerful model for describing the activity of allosterically regulated enzymes[18–22], which assumes that allosteric enzymes exist in two or more conformations with different kinetic properties. Both the binding of substrates and allosteric regulators can modify the kinetic properties of an MWC enzyme by stabilizing one conformation over another. We use statistical approaches–regularization and cross-validation–to identify the simplest MWC kinetic rate equation that can adequately describe the available data, allowing us to avoid overfitting and parameter identifiability issues common for fitting complex equations to finite data[30]. The MWC equation describing the rate of PFKP in the presence of substrates and regulators is shown in Figure 1D. We described the rest of the glycolytic enzymes using standard kinetic rate equations derived from quasi-steady-state or rapid equilibrium approximations. For the eight enzymes and transporters with one substrate and one product, we used reversible Michaelis-Menten equations (see equation describing the rate of glucose-6-phosphate isomerase (GPI) in Figure 1D) and estimated their kinetic constants by averaging the values from at least three publications per constant to verify their consistency and accuracy. For the three remaining enzymes with more than one substrate or product (i.e., aldolase (ALDO), phosphoglycerate kinase (PGK), and lactate dehydrogenase (LDH)), we fitted their more complex rate equations to manually curated *in vitro* kinetics datasets containing 350-700 data points per enzyme. We chose to fit data from ALDO, PGK, and LDH instead of averaging over published kinetic constants because many publications describing ALDO, PGK, and LDH

activity used different rate equations, and hence the published kinetic constants are not directly comparable. A comprehensive description of the derivations and fitting of all the enzyme rate equations is reported in Notes S1-S3, and we hope this compendium of information will serve as a great resource for future investigations of these enzymes.

Our model uses several assumptions that must be considered when interpreting its predictions. First, the model assumes that the activity of enzymes in living cells is accurately described using *in vitro* activity of purified enzymes. Second, the model assumes that enzymes and metabolites are well-mixed in the cytosol of the cell. Third, the model only describes the effect of regulators that are included in the model. Fourth, the model assumes that other pathways (e.g., pentose phosphate pathway, mitochondrial pyruvate consumption) do not affect the concentration of glycolytic intermediates or glycolytic reaction rates. Deviations between the model and experimental observations could be due to any number of these assumptions not being valid under the conditions of a particular experiment. On the other hand, if the model accurately predicts experimental data, it suggests that these assumptions are valid under given experimental conditions.

Glycolysis model robustly performs the key tasks of glycolytic ATP homeostasis

After constructing the model, we first performed simulations to determine whether the glycolysis model can perform the key tasks of ATP homeostasis (Figure 1A). To model ATP consumption by intracellular processes, we added an ATPase enzyme to the model. Glycolysis cannot proceed faster than the maximal rate of its slowest enzyme, so we report all ATPase rates in terms of the percent of the maximal activity of HK1–the slowest enzyme of glycolysis based on maximal rates determined by the product of intracellular enzyme concentration and V_{max}. First, we found that the glycolysis model produces as much ATP as is consumed and rapidly adjusts to stepwise increase or decrease of ATP consumption, as is necessary under physiological conditions (Figure 2A). Second, we showed that > 90% of the adenine nucleotide pool was maintained as ATP and ATP concentrations varied by < 10% when we introduced physiologically relevant 2-fold stepwise changes in ATP consumption rate (Figure 2B). Thus, as has been observed *in vivo*[24,25], our model maintains nearly constant ATP levels in response to large changes in ATP turnover. Finally, we showed that the glycolysis model maintains ATP, ADP and inorganic phosphate concentrations such that the ATP hydrolysis reaction is ~10^9-fold away from equilibrium, equivalent to ~21 k_BT (Figure 2C), which is similar to what has been measured in cells[23].

We next tested whether the ability of the model to perform the relevant tasks of ATP homeostasis is robust to changes in glycolytic enzyme levels that are variable from cell to cell and to uncertainty in our estimates of biochemical constants, cofactor pool concentrations, and initial concentrations of intermediates. To demonstrate robustness, we reran our model 10,000 times using random values of enzyme levels spanning a 9-fold range around the default model estimate (Figure S1A-C) as well as random values of model parameters (Figure S1D-F) and initial conditions (Figure S1G-K) based on their respective experimental mean and standard error. Our simulations showed that 95% confidence of model prediction is largely independent of glycolytic enzyme levels, model parameters and initial conditions as would be expected for the model that captures the robust endogenous regulatory circuit of glycolytic ATP homeostasis.

Several kinetics models of glycolysis have been reported over the past several decades[4–17], yet none of these reports investigated whether those models can recapitulate the tasks of ATP

homeostasis except for matching ATP supply and demand. To perform a head-to-head comparison, we implemented a recent mechanistic model of yeast glycolysis that we will refer to as the van Heerden model[15]. Upon repeating the simulations in Figure 2A-C, the van Heerden model produced ATP in response to ATP consumption and maintained the high energy of ATP hydrolysis (Figure S2A, C). However, the van Heerden model could only support ATP levels that are orders of magnitude lower than our model and exhibited 10-fold volatility in ATP concentration in response to 2-fold stepwise changes in ATP demand compared to the <10% change exhibited by our model under the same conditions (Figure S2B). Thus, the van Heerden model does not capture the regulation that is needed to maintain most of the adenine nucleotide pool in the form of ATP and ensure stable ATP levels in response to changes in ATPase rate as observed *in vivo*[24,25]. Furthermore, van Heerden and other published models simulated glycolysis at constant levels of inorganic phosphate, which, as we demonstrate below, cannot be used to study allosteric regulation as its function is to maintain inorganic phosphate levels.

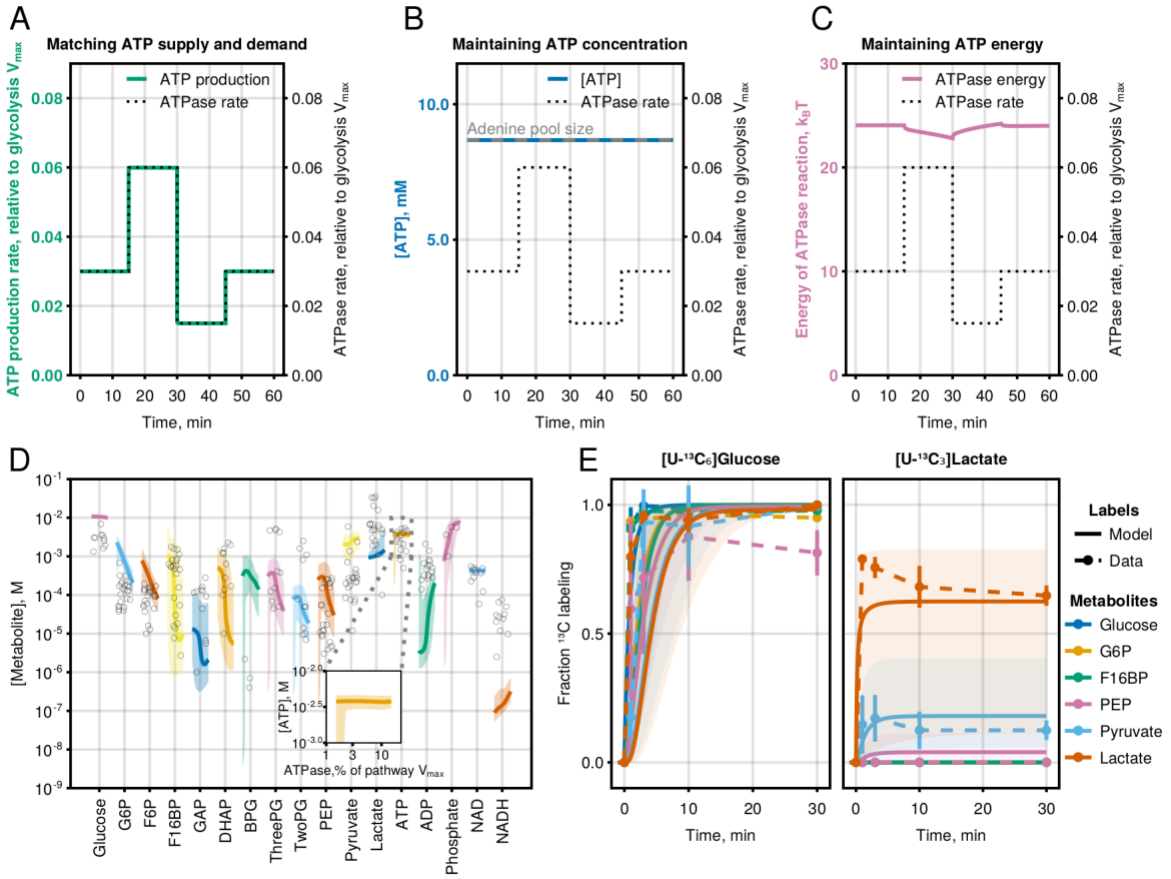

Figure 2. Glycolysis model recapitulates glycolysis activity observed in live cells. (A-C) Model simulations showing changes in (A) ATP production rate, (B) ATP concentration (total adenine pool size is labeled with dashed grey line), and (C) energy released during ATP hydrolysis in response to a 2-fold stepwise increase or decrease of ATPase rate. Starting ATPase rate = 3% of glycolysis Vmax was used for (A-C). ATPase energy is calculated as a natural logarithm of the disequilibrium ratio for the ATPase reaction (i.e., mass-action ratio divided by the equilibrium constant) for the ATPase reaction. **(D)** Comparison of measured glycolytic intermediate concentrations and model predictions. Black empty circles indicate the

experimentally determined metabolite concentrations. Colored lines with ribbon indicate the median and 95% CI of model predictions. The shift from left to right for each metabolite level prediction corresponds to increasing rates of ATP consumption from 1.5% to 15% of glycolysis Vmax (see inset for ATP) to highlight the effect of ATP consumption on metabolite concentrations. Model predictions represent the concentration of bound + free metabolites in the whole cell analogous to whole cell metabolite measurements, while in all other figures, model output is the free concentration of metabolites in the cytosol. (**E)** Comparison of [U-$^{13}C_6$]Glucose and [U-$^{13}C_3$]Lactate tracing data with model prediction. Circles connected with dotted lines are data, and lines with ribbons are model predictions. ATPase rate = 15% of glycolysis Vmax was chosen for model predictions. Model predictions in (D) and (E) are from simulations where bootstrapped model parameter combinations could match ATP supply and demand, which were > 92% and >95% of simulations for (D) and (E), respectively. See also Figure S1 and Figure S2. Julia code to reproduce this figure is available at https://github.com/DenisTitovLab/Glycolysis.jl.

Glycolysis model accurately predicts metabolite levels and fluxes in living cells

Given that our model could reproduce the wealth of measurements of each glycolytic enzyme *in vitro* and perform the key tasks of ATP homeostasis, we set out to determine whether the model could predict the results of experiments in live cells.

To directly compare model predictions to data from bulk cell measurements, we adjusted the model output in two ways. First, we corrected for the fact that metabolites inside the cell can be either bound to proteins or free in solution. Concentration terms in enzyme rate equations refer specifically to free metabolite concentrations, so we converted free metabolite concentrations calculated by the model to free + bound concentrations using the binding constants of metabolites to enzymes that can be derived from kinetic rate equations. An important caveat of this calculation is that we only consider glycolytic enzymes, which neglects the binding of glycolytic intermediates to other enzymes. Second, we corrected for the fact that glycolytic enzymes are localized to the cytosol and that a significant fraction of the cytosol is occupied by macromolecules and not water. The latter effectively means that all the enzymes and metabolites of glycolysis are concentrated in a volume that is ~50% of cellular volume, assuming ~70% of the cell is cytosol[31] and ~70% of the cytosol is water[32,33].

We first compared the model predictions of bound and free metabolites with absolute intracellular metabolite concentrations. Mammalian cells can synthesize ATP using aerobic glycolysis or mitochondrial respiration. Since our model describes ATP homeostasis by aerobic glycolysis, the most appropriate comparison would be with cells that don't have a functional respiratory chain like red blood cells or cells treated with respiratory inhibitors. We have measured absolute concentrations of glycolytic intermediates in four cell lines using liquid chromatography-mass spectrometry (LC-MS). To ensure the most comprehensive comparison of our model prediction to data, we also used published data from red blood cells as well as cancer cell lines. We found that 95% confidence intervals of model predictions of most of the 18 glycolytic intermediate concentrations at a range of ATP consumption rates overlapped within experimental values (Figure 2D). The latter result is notable given that our model contains no direct information about intracellular metabolite levels and is free to predict concentrations from 0 to +∞ and even below zero if the model is implemented incorrectly.

Next, we compared [U-$^{13}C_6$]Glucose and [U-$^{13}C_3$]Lactate labeling kinetics predicted by the model to cellular measurements. For these measurements, we exchanged normal media for media containing [U-$^{13}C_6$]Glucose or [U-$^{13}C_3$]Lactate and then lysed cells at different time intervals to estimate the rate and fraction at which ^{13}C from [U-$^{13}C_6$]Glucose or [U-$^{13}C_3$]Lactate is incorporated into glycolytic intermediates. We observed that the glycolysis model accurately recapitulated the kinetics of the ^{13}C labeling fraction of intermediates after switching to [U-$^{13}C_6$]Glucose or [U-$^{13}C_3$]Lactate-containing media (Figure 2E). It is particularly noteworthy that the model can recapitulate labeling from [U-$^{13}C_3$]Lactate where only intracellular lactate and pyruvate are labeled, as this indicates that our model can accurately capture the extent of *reversibility* of glycolysis reactions.

Allosteric regulation is required to robustly maintain high ATP levels while mass action performs other tasks of glycolytic ATP homeostasis

Having established that the model recapitulates the key tasks of ATP homeostasis and accurately predicts measurements in live cells, we next used the model to investigate the function of allosteric regulation and mass action in maintaining ATP homeostasis.

We first removed all allosteric regulators to see if the resulting simple pathway remained functional. Specifically, we made the HK1, PFKP, GAPDH, and PKM2 enzymes behave as Michaelis-Menten-like enzymes with kinetic parameters corresponding to their faster (active) MWC conformation. We also removed competitive inhibition of the HK catalytic site by G6P, which proceeds through standard non-allosteric competitive inhibition. Surprisingly, the model without any regulation was fully capable of matching the rates of ATP supply and demand and maintaining the high energy of ATP hydrolysis (Figure 3A, C, D, F). However, without allosteric regulation, we observed a complete breakdown of high ATP level maintenance, where ATP levels were >100-fold lower, and a small 2-fold increase and decrease of ATPase rate led to an almost 10-fold change in ATP concentration compared to <10% change in ATP concentration for the complete model (Figure 3B, E). Systematic investigation of steady-state model behavior at different ATPase rates showed that the model without regulation is not capable of maintaining most of the adenine nucleotide pool in the form of ATP levels at any ATPase value, which is contrary to the constant ATP concentrations observed both *in vivo*[24,25] and in the complete model (Figure 3E, vertical drop-off indicates where the ATP production by the model cannot match ATPase rate anymore). We also note that our model without allosteric regulation behaved similarly to the van Heerden model with allosteric regulation, indicating that the latter does not capture the role of allostery in regulating glycolysis (Figure 3B and Figure S2B).

We performed a global sensitivity analysis of our model to systematically explore the role of all model parameters in matching ATP supply and demand, maintaining high ATP levels, and maintaining high energy of ATP hydrolysis. The goal of global sensitivity analysis is to calculate the contribution of each model parameter and parameter interactions to the variance of a specific model output[34]. We used the average ratio of ATP production to ATP consumption, average ratio of ATP concentration to adenine pool size, and average energy of ATP hydrolysis at a log range of ATPase values as proxies for the model's ability to match ATP supply and demand, maintain high ATP levels and maintain high energy of ATP hydrolysis, respectively. We first estimated the variance of the proxies by randomly varying each parameter of the model independently from a uniform distribution spanning a 9-fold range from 3 times lower to 3 times higher than the default model parameter values. The coefficient of variation (CV) for all three AUC proxies was

in the narrow range of 0.13-0.34 in response to random changes in parameter values spanning 9-fold, indicating that these are robust properties of the model (Figure S3A-C). The coefficient of variation for maintaining high ATP levels was about 3-fold larger than for maintaining high energy of ATP hydrolysis or for matching ATP supply and demand, suggesting that the former is more sensitive to changes in specific model parameters. Two metrics for each model parameter are typically reported for variance-based global sensitivity analysis. The first-order effect sensitivity index S_1 reports the fraction of variance of the model output that will be removed if the corresponding parameter is fixed. The total-order sensitivity index S_T reports the variance that will be left if values of all but the corresponding parameter are fixed. Larger values of S_1 and S_T indicate that a given parameter is important. Global sensitivity analysis showed that the K_M and V_{max} for HK1 had the highest S_1 and S_T for all three model outputs, as expected since HK1 is the slowest enzyme in the pathway. The second largest S_1 and S_T for maintaining high ATP levels were given by the kinetic parameters that control allosteric regulation of HK1 and PFKP, while the second largest S_1 and S_T for matching ATP supply and demand and for maintaining high energy of ATP hydrolysis were simply the K_M and V_{max} of PFKP or the second slowest enzyme in the model after HK1 based on the product of enzyme concentration and V_{max} (Figure S3D-F).

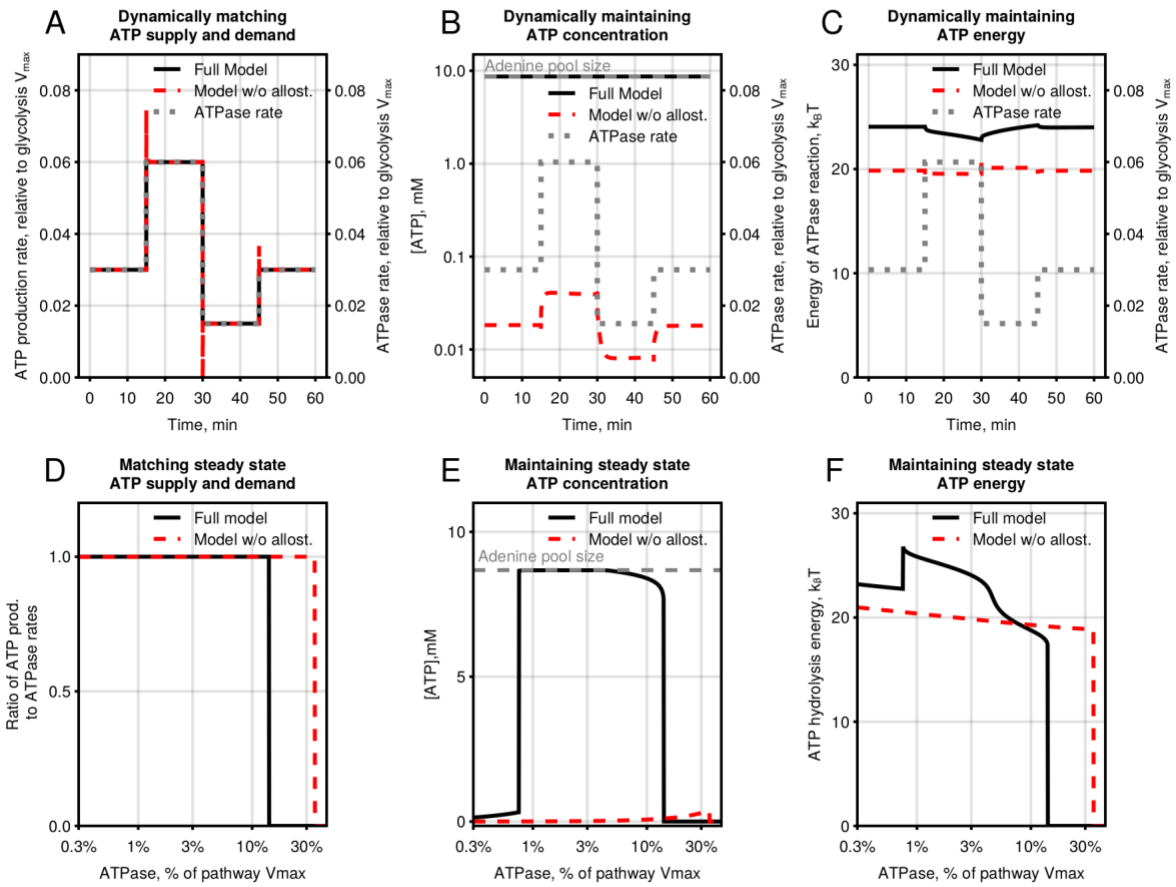

Figure 3. Allosteric regulation is required for robust maintenance of high and stable ATP levels. (A-C) Model simulations showing the effect of removal of allosteric regulation of HK1, PFKP, GAPDH, and PKM2 on dynamic changes in (A) ATP production rate, (B) ATP concentration, and (C) energy released during ATP hydrolysis [Left axis] in response to a 2-fold

stepwise increase or decrease of ATPase rate [Right axis]. Starting ATPase rate = 3% of glycolysis Vmax was used for (A-C). Energy is calculated as a natural logarithm of the disequilibrium ratio (i.e., mass-action ratio divided by the equilibrium constant) for the ATPase reaction. **(D-F)** Model simulations showing the effect of removal of allosteric regulation of HK1, PFKP, GAPDH, and PKM2 on steady-state (D) ATP production rate, (E) ATP concentration, and (F) energy released during ATP hydrolysis at a 100-fold range of ATPase rates [Right axis]. Dashed grey lines in B and E indicate the total adenine pool size (i.e., ATP + ADP + AMP). See also Figure S3. Julia code to reproduce this figure is available at https://github.com/DenisTitovLab/Glycolysis.jl.

Redundant allosteric regulators of HK1 and PFKP maintain high and stable ATP levels

To identify the role of specific allosteric regulators in maintaining high and stable ATP levels, we computationally dissected the pathway by removing the allosteric regulation of enzymes one by one (Figure 4A-C). Specifically, we removed a particular allosteric regulator from the relevant kinetic rate equations by setting its binding constant for both active and inactive MWC conformation to ∞ and setting constant L that determines the ratio of inactive to active MWC conformations to zero. We found that the allosteric regulation of HK1 and PFKP is responsible for maintaining high and stable ATP levels, whereas removing the allosteric regulation of GAPDH and PKM2 had no discernable effect (Figure 4A-D). Digging further, we removed each of the allosteric activators and inhibitors of HK1 and PFKP and found that these regulators work together to ensure the robust maintenance of ATP levels, so removing any single regulator only led to a partial loss of ATP maintenance capacity. In general, removing inhibitors–G6P for HK1 and ATP for PFKP–led to worse ATP maintenance at low ATPase rates (Figure 4E-H), while removing activators–phosphate for both HK1 and PFKP and ADP for PFKP–led to poorer ATP maintenance at high ATPase rates (Figure 4K-N).

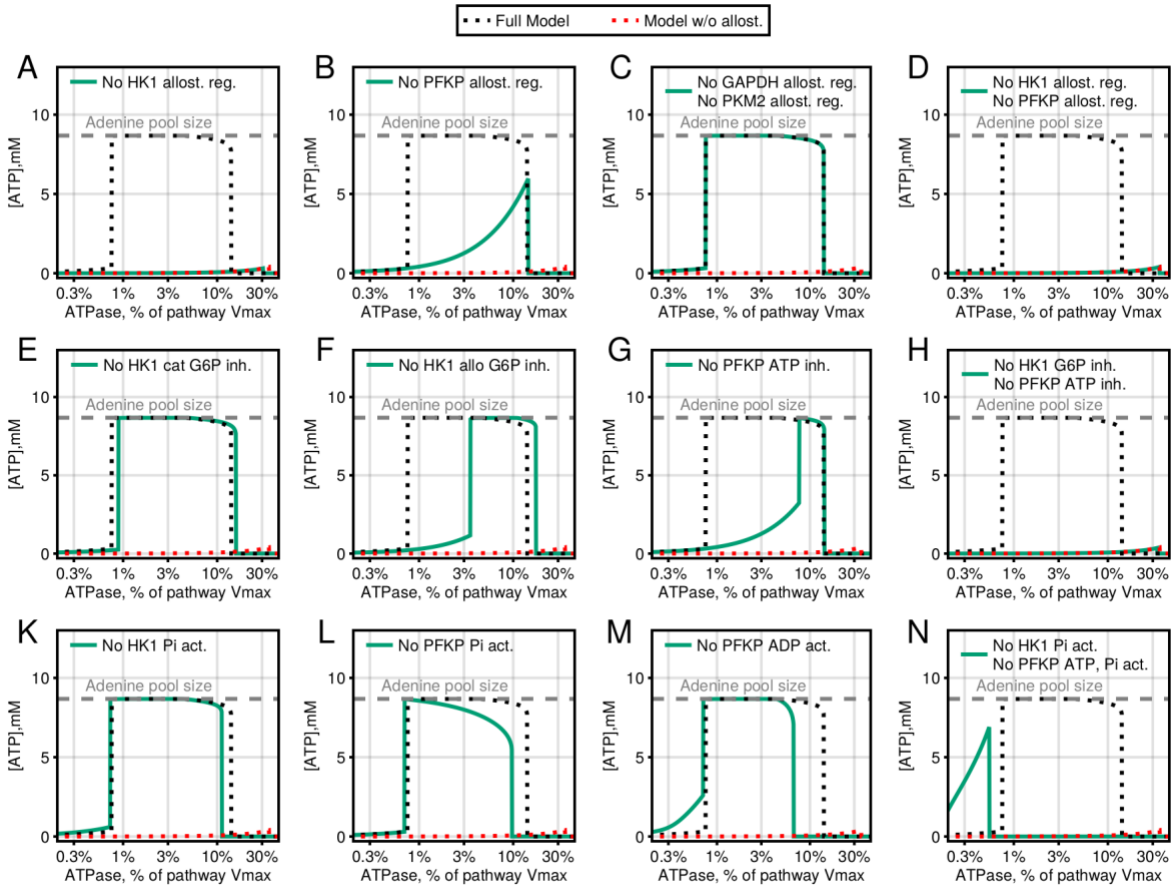

Figure 4. Redundant allosteric regulators of HK1 and PFKP are required to maintain high and stable ATP levels. (A-N) Steady-state ATP concentrations at a range of ATPase rates of the glycolysis model with and without (A-D) allosteric regulation of enzymes, (E-H) allosteric inhibitors, and (K-N) allosteric activators. Dashed grey line indicates the total adenine pool size (i.e., ATP + ADP + AMP). Julia code to reproduce this figure is available at https://github.com/DenisTitovLab/Glycolysis.jl.

Allosteric regulation maintains ATP levels by preventing phosphate disappearance due to an imbalance between lower and upper glycolysis

We next investigated the mechanism that allows allosteric regulators of HK1 and PFKP to maintain high and stable ATP levels such that most of the adenine pool is in the form of ATP. Here, we first provide an intuitive explanation of the mechanism and then show simulations supporting it. ATP levels are bounded by the size of the adenine pool from the top and the ability of glycolysis to convert most of the ADP into ATP from the bottom. The latter leads to near-constant ATP levels under any conditions where glycolysis can convert most of the ADP into ATP. Glycolysis has an unusual organization where upper glycolysis–enzymes from HK1 to ALDO–consumes ATP, while lower glycolysis–enzymes from GAPDH to LDH–consumes inorganic phosphate and produces ATP. Such an organization can lead to a futile cycle where lower glycolysis converts inorganic phosphate to ATP that is used by upper glycolysis to produce phosphorylated glycolytic intermediates upstream of GAPDH (Figure 5A). This futile

cycle is thermodynamically favorable, resulting in the trapping of inorganic phosphate in phosphorylated intermediates of upper glycolysis. Low levels of inorganic phosphate inhibit GAPDH reaction, which leads to a lower concentration of substrates of ATP-producing enzymes of lower glycolysis–BPG for PGK and PEP for PKM2–requiring higher levels of ADP to drive PGK and PKM2 forward through mass action. The function of allosteric regulation of HK1 and PFKP is to prevent this futile cycle by matching the rates of HK and PFK to half of the ATPase rate.

We performed simulations to support the mechanism shown in Figure 5A. Removal of allostery leads to upregulation of all the metabolites before the GAPDH step except for phosphate and downregulation of all the metabolites after the GAPDH step, suggesting that low phosphate blocks the GAPDH step in the absence of allostery (Figure 5B). We next showed that keeping the levels of inorganic phosphate constant abrogated the requirement for allostery to maintain ATP levels, showing that allostery works by maintaining high phosphate levels (Figure 5C). Similarly, a large decrease in equilibrium constants of HK1, PFKP, GAPDH, and PGK that makes the futile cycle of phosphate less thermodynamically favorable also removed the requirement for allostery (Figure 5D). Finally, substituting HK1 and PFKP rate equations for one-half of the ATPase rate abolished the requirement for allostery, suggesting that the latter works by tying HK1 and PFKP rates to the ATPase rate, which makes the futile cycle impossible at a range of ATPase rates.

We note that the previously published models of glycolysis were investigated under conditions with constant phosphate levels, which made it impossible to identify the role of allosteric regulation.

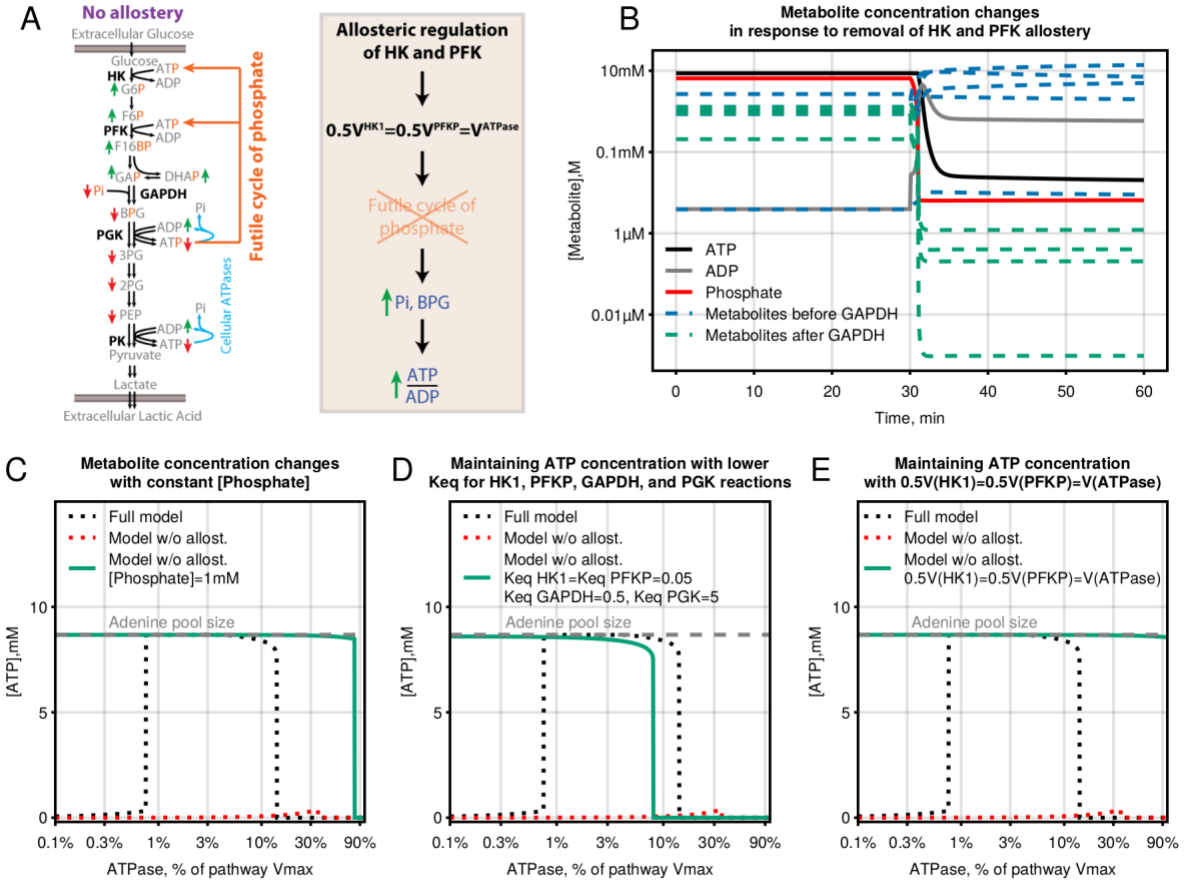

Figure 5. Allosteric regulation maintains ATP levels by preventing phosphate disappearance due to an imbalance between lower and upper glycolysis. A) Schematic of the mechanism of ATP maintenance by allosteric regulation of HK1 and PFKP. **B)** Metabolite concentration changes upon an instantaneous removal of allostery at 30 min. ATPase rate = 3% of glycolysis Vmax was used for model predictions. **C-E)** Steady-state ATP concentrations at a range of ATPase rate of the glycolysis model with (C) constant [Phosphate] = 1mM, (D) K_{eq}^{HK1}=0.05 (default model K_{eq}^{HK1}=2,700), K_{eq}^{PFKP}=0.05 (default model K_{eq}^{PFKP}=760), K_{eq}^{GAPDH}=0.5 (default model K_{eq}^{GAPDH}=16), and K_{eq}^{PGK}=5 (default model K_{eq}^{PGK}=2000), (E) Rate equations for HK1 and PFKP substituted for one-half of ATPase rate equations. Julia code to reproduce this figure is available at https://github.com/DenisTitovLab/Glycolysis.jl.

Excess activity of enzymes relative to HK and PFK is required to maintain high ATP levels

We next explored whether attributes of the glycolysis pathway other than allosteric regulation of HK1 and PFKP might be required for maintaining most of the adenine pool in the form of ATP. We noted that most glycolytic enzymes are expressed in large excess compared to HK1, PFKP, glucose transporter, and lactate transporter (Figure 6A). In fact, most enzymes are present at 10-100-fold higher concentrations than is required for the maximal activity of the glycolysis pathway limited by HK1 (Figure 6B). Allosteric inhibition of HK1 and PFKP by G6P and ATP, respectfully, further increases the differential between the activity of HK1 and PFKP and the rest

of the enzymes. We used our model to understand the rationale for this seemingly wasteful enzyme expression pattern, given that production of excess glycolytic enzymes represents a significant investment of resources by the cell with enzymes like GAPDH and PKM2 approaching 1% of the proteome. Increasing the concentrations of all enzymes from TPI to LDH by 10-fold more than default model values led to no changes in ATP maintenance (Figure 6C). Decreasing the expression of these enzymes by 10- and 50-fold led to a small decrease and complete collapse of ATP level maintenance, respectively (Figure 6C). Low expression of HK1 and PFKP contributes to limiting the occurrence of the futile cycle of inorganic phosphate, as the latter requires the excess activity of HK1 and PFKP. Thus, the relative expression of glycolytic enzymes works together with allosteric regulation to maintain high and stable ATP levels. Our results provide a mechanistic explanation for the seemingly wasteful large excess expression of enzyme of glycolysis relative to HK and PFK.

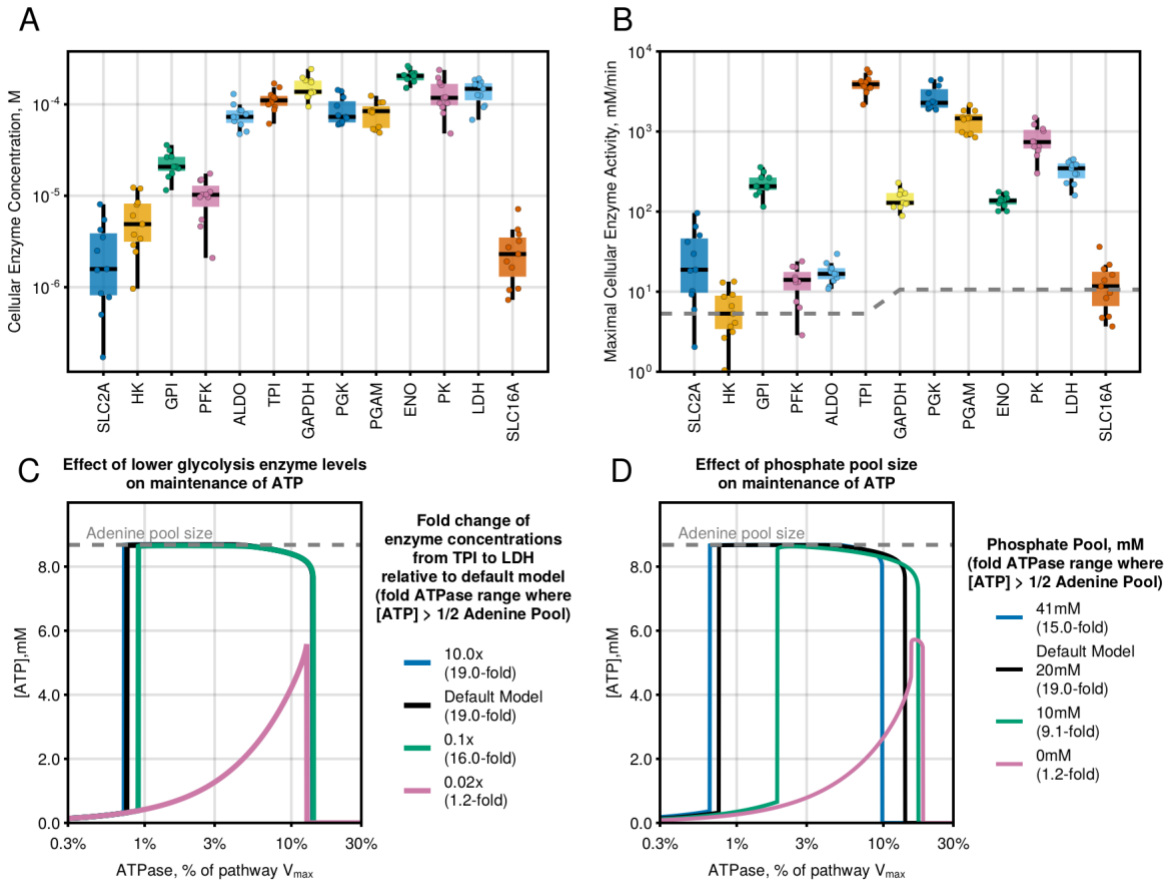

Figure 6. Excess activity of glycolytic enzymes relative to HK and PFK and large inorganic phosphate pool improve the maintenance of high and stable ATP levels. (A) Cytosolic enzyme concentrations in the model estimated from proteomics data. **(B)** Maximal cytosolic enzyme activities in the model. Activities of all isoforms for each enzyme were summed together. The dashed line indicates the activity for each enzyme that matches the maximal activity of the slowest enzyme HK1 corrected for the stoichiometry of reactions. **(C)** Steady-state ATP concentrations at a range of ATPase rates of the glycolysis model with scaled concentrations of lower glycolysis enzymes TPI, GAPDH, PGK, PGAM, ENO, PK, and LDH. **(D)** Steady-state ATP concentrations at a range of ATPase rates of the glycolysis model with

different levels of phosphate pool size. The dashed grey line indicates the total adenine pool size (i.e., ATP + ADP + AMP). The fold range of ATPase rates where the model can maintain >50% of the adenine nucleotide pool as ATP is highlighted in legend. See also Figure S4. Julia code to reproduce this figure is available at https://github.com/DenisTitovLab/Glycolysis.jl.

Large inorganic phosphate pool size improves the maintenance of high ATP levels

We next investigated the role of cofactor pool sizes in maintaining ATP levels. The three cofactor pools that are included in our model are the adenine nucleotide pool (i.e., ATP+ADP+AMP), NAD(H) pool (i.e., NAD^+ + NADH), and the inorganic phosphate pool (i.e., the sum of all phosphate atoms except for ATP, ADP, and AMP). For the inorganic phosphate pool, we did not consider phosphate that is part of the adenine nucleotide pool so we could separate the effect of inorganic phosphate and adenine nucleotide pool sizes. We performed a global sensitivity analysis by varying cofactor pool sizes over a 9-fold range centered around the default values of each pool size. Based on S_1 and S_T sensitivity indexes, phosphate pool size had the largest effect on the model's ability to maintain high ATP levels, while adenine pool size had a smaller effect and NAD(H) pool size had little to no effect (Figure S4A-B).

To confirm the effect of the phosphate pool on maintaining ATP levels, we simulated an increase or decrease of the level of all phosphorylated metabolites except ATP, ADP, and AMP. The simulations confirmed that changes in the phosphate pool led to a large change in the model's ability to maintain high and stable ATP levels with the near-complete collapse of ATP maintenance in the absence of the inorganic phosphate pool (Figure 6D). It is noteworthy that the default phosphate pool size used in our model that is estimated from metabolomics measurements is close to optimal as both decreasing and increasing phosphate pool led to the deterioration of the model's ability to maintain high and stable ATP levels. The effect of phosphate pool size is driven by the concentration of inorganic phosphate, which, if kept constant, abrogates the effect of a changing phosphate pool (Figure S4C). Cells have complex regulatory machinery to regulate intracellular phosphate pools, and it is likely that this machinery–that is not part of our model–works together with allosteric regulation and relative expression of glycolytic enzymes to maintain high and stable ATP levels by preventing phosphate concentration from dropping below the minimal level sufficient for GAPDH activity.

Finally, we confirmed the effect of adenine pool size on ATP level maintenance using simulations where we increased or decreased the combined level of ATP, ADP, and AMP (Figure S4D). As for the inorganic phosphate pool, the size of the adenine nucleotide pool that we estimated from metabolomics data and used in our model was optimal for the ability of the model to support high and stable ATP levels. The observations that both inorganic phosphate pool and adenine nucleotide pool sizes estimated from metabolomics data are optimal for the ability of our model to support high and stable ATP levels provide additional evidence of the capacity of our model to describe the physiologically relevant behavior of glycolysis.

Simple two-enzyme model of glycolysis recapitulates key tasks of ATP homeostasis

To gain a better understanding of the tasks of ATP homeostasis that can be achieved by mass action alone and which might require allostery, we used a simplified model of glycolysis,

referred to as the two-enzyme model, containing ATP-consuming Enzyme 1, ATP-producing Enzyme 2 and ATPase (Figure 7A):

Enzyme 1: $G + ATP \rightleftarrows X + ADP$

Enzyme 2: $X + Phosphate + 2ADP \rightleftarrows L + 2ATP$

ATPase: $ATP + H_2O \rightleftarrows Phosphate + ADP$

Note that metabolite X, which represents all intracellular glycolysis intermediates, is phosphorylated to conserve phosphate in each step.

To simplify the analysis, we described the kinetics of Enzyme 2 and ATPase using modified reversible Michaelis-Menten equations with the assumption that Enzyme 2 and ATPase are saturated with substrates[35]. For Enzyme 1, we used the MWC model to add allosteric activation by ADP and phosphate and allosteric inhibition by ATP. We assumed that ADP and Phosphate only bind to active conformation, ATP only binds to inactive conformation, inactive MWC conformation of Enzyme 1 is catalytically inactive, and regulatory and catalytic sites are saturated with regulators and substrates. The resulting two-enzyme model is described using the following rate equations (see STAR Methods for additional detail on the derivation of the two-enzyme model rate equations):

$$V^{Enz\,1} = \frac{V_{max}^{Enz\,1}}{\left(1 + \left(\frac{L[ATP]}{[ADP][Phosphate]}\right)^n\right)} \left(1 - \frac{1}{K_{eq}^{Enz\,1}} \frac{[X][ADP]}{[G][ATP]}\right) \quad (1)$$

$$V^{Enz\,2} = V_{max}^{Enz\,2} \left(1 - \frac{1}{K_{eq}^{Enz\,2}} \frac{[L][ATP]^2}{[X][ADP]^2[Phosphate]}\right) \quad (2)$$

$$V^{ATPase} = V_{max}^{ATPase} \left(1 - \frac{1}{K_{eq}^{ATPase}} \frac{[ADP][Phosphate]}{[ATP]}\right) \quad (3)$$

Two-enzyme model reduces the number of parameters from >150 for the full model to only 8 for the two-enzyme model, including only two parameters–*L* and *n*–responsible for allosteric regulation.

We searched for values of the parameters of the two-enzyme model that would support the three tasks of ATP homeostasis. We performed 10,000 simulations using random values of two-enzyme model parameters with or without allostery to see which combinations of parameters can support the three tasks at a range of ATPase values (Figure 7B-D). All the parameter combinations could match ATP supply and demand and maintain the energy of ATP hydrolysis >9 k_BT in the presence or absence of allostery, confirming that mass action alone is fully capable of supporting these two tasks (Figure 7B, D). By contrast, only a small fraction of simulations in the presence of allostery could maintain ATP > ADP (Figure 7C). As for the full model (Figure 5C), keeping the phosphate level constant abrogated the requirement for allostery, confirming that maintenance of ATP > ADP is a non-trivial task that requires allostery to maintain high levels of inorganic phosphate (Figure 7C, F).

We next performed simulations of the two-enzyme model with the parameters that supported the highest level of ATP at a range of ATPase values. We found that removing allostery by setting *L=0* completely abrogated while decreasing *n* from 4 to 1 attenuated ATP level maintenance (Figure 7E). The effect of *n* suggests an important physiological role for the oligomeric structure of allosteric enzymes like PFKP. As for the full model (Figure 5D), decreasing the equilibrium constant of Enzyme 1 abrogated the requirement for allostery as the futile cycle of phosphate is less thermodynamically favorable under these conditions. It is noteworthy that the two-enzyme

model without allostery could maintain ATP>ADP under one exact condition when the $V_{max}^{Enz\,1} = V_{max}^{ATPase}$ (red line in Figure 7E at 100% ATPase rate)–a condition when the futile cycle of phosphate would be impossible as all of the Enzyme 1 capacity is used to satisfy ATP demand with no spare capacity to catalyze the futile cycle reactions. As for the full model (Figure 5E), making $V_{max}^{Enz\,1} = V_{max}^{ATPase}$ abrogated the requirement for allostery (Figure 7H). The latter again highlights that allostery prevents the futile cycle of phosphate by adjusting $V_{max}^{Enz\,1}$ such that $V_{max}^{Enz\,1} = V_{max}^{ATPase}$ at a large range of V_{max}^{ATPase}.

Overall, our analysis of the two-enzyme model confirms the results of the full model that allosteric regulation is required to maintain high and stable ATP levels, while mass action alone is sufficient to match ATP supply and demand and to ensure that ATP hydrolysis generates energy.

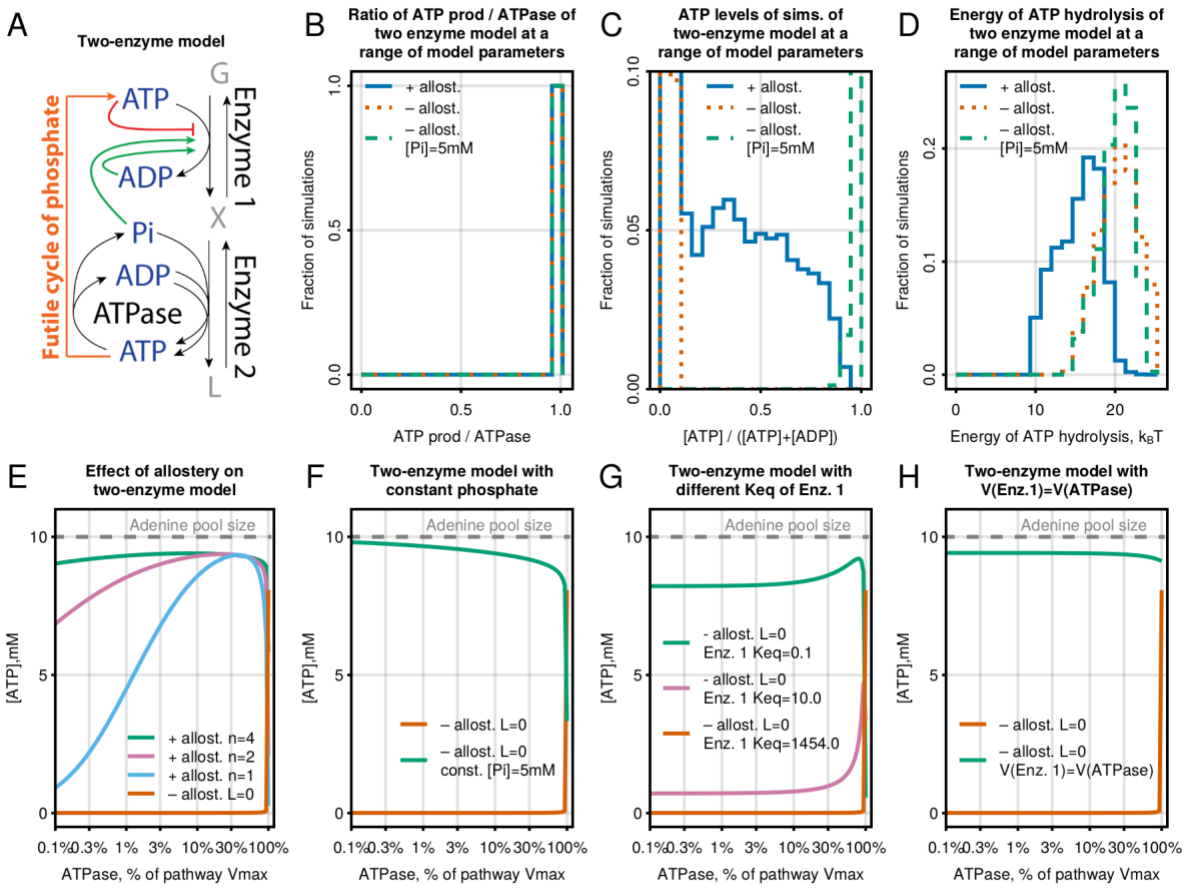

Figure 7. Simple coarse-grained model of glycolysis recapitulates key tasks of ATP homeostasis. (A) Schematic of a simplified glycolysis-like pathway containing two enzymes with feedback on Enzyme 1. **(B-D)** Simulation of the two-enzyme model at random values of the 6 parameters controlling this pathway showing the ability to (B) match ATP supply and demand, (C) maintain ATP levels in relation to ATP+ADP, and (D) generate energy from ATP hydrolysis. **(E-H)** Simulation of the two-enzyme with best-performing parameters from (C) showing maintenance of steady-state ATP levels at a 100x range of ATPase rates with (E) presence and absence of allostery, (D) constant inorganic phosphate levels, (G) different K_{eq} for

Enzyme 1 and (H) rate of Enzyme 1 equal to rate of ATPase. Julia code to reproduce this figure is available at https://github.com/DenisTitovLab/Glycolysis.jl.

DISCUSSION

Metabolism is an emergent property of the activities of many individual enzymes. Regulation of metabolic pathways has historically been studied by rigorously characterizing enzymes in isolation. While this has greatly advanced our understanding of the key regulatory enzymes, many emergent functions of metabolism are difficult to understand by focusing on one enzyme at a time. Approaches that combine theory and experiments are needed to understand how the activities of individual enzymes give rise to metabolic homeostasis. Here, we developed a data-driven model of glycolysis that uses enzyme rate equations to describe the activity of the whole glycolysis pathway and applied it to identify the mechanisms that allow glycolysis to perform the three key tasks of ATP homeostasis (Figure 1A). The resulting model can quantitatively predict the output of glycolysis in response to perturbations of metabolite levels and kinetic properties of enzymes, such as changes in media conditions, enzyme expression, drug treatment, and expression of enzyme mutants. In the discussion section, we highlight the ability of our model to explain a wide range of experimental observations.

The textbook view of glycolysis regulation has been that allostery controls the rate of glycolytic ATP production. We show that mass action alone is sufficient for regulating glycolytic ATP production rate to match ATP supply and demand, and for maintaining ATP, ADP, and inorganic phosphate concentrations such that ATP hydrolysis can generate energy (Figure 3, Figure 7). The key role of mass action is perhaps not surprising given the stoichiometry of the net reaction of glycolysis:

$$Glucose + 2ADP + 2Phosphate \rightleftarrows 2Lactate + 2ATP \qquad (4)$$

An increase in ATP consumption rate would push this reaction forward by mass action due to a decrease in the product (ATP) and an increase in substrate levels (ADP and phosphate) resulting from ATP hydrolysis. Likewise, the ATP hydrolysis reaction will be maintained far from equilibrium as long as the net reaction of glycolysis is thermodynamically favorable and can proceed faster than ATP hydrolysis. Thus, the ability of glycolysis to control the rate of ATP production to match ATP supply and demand, and maintain ATP hydrolysis far from equilibrium are built-in properties of the glycolytic pathway that are largely independent of specific enzyme kinetic mechanisms.

If mass action controls the rate of ATP production, what is, then, the function of allosteric regulation? We showed that the allosteric regulation of HK1 and PFKP by G6P, ATP, ADP, and Pi is required to maintain most of the adenine nucleotide pool in the form of ATP without affecting the rate of ATP production (Figure 3A-C). Maintenance of ATP > ADP + AMP is a difficult task due to the structure of the glycolysis pathway where a futile cycle can be formed between ATP-producing enzyme of lower glycolysis PGK and ATP-consuming enzymes of upper glycolysis HK1 and PFKP that leads to the trapping of phosphate in upper glycolysis intermediates and the resulting disappearance of free inorganic phosphate and phosphorylated intermediates of lower glycolysis (Figure 5A). Low levels of phosphorylated intermediates of lower glycolysis necessitate higher ADP levels to drive ATP-producing reactions of lower glycolysis through mass action. As a consequence of this futile cycle, ADP levels become higher than ATP levels but the rate of ATP production is not affected as high ADP compensates for low

levels of phosphorylated intermediates of lower glycolysis and the energy of ATP hydrolysis is not affected either as both ATP/ADP ratio and phosphate levels drop, preserving the ATP / ADP • Phosphate ratio (Figure 3A-C). We showed that allosteric regulation of HK1 and PFKP prevents this futile cycle by tying the rates of HK1 and PFKP to one-half of the ATPase rate such that there's no spare catalytic activity of HK1 and PFKP to catalyze the futile cycle reactions.

One of the remarkable properties of ATP homeostasis is the maintenance of stable ATP levels in response to up to 10-fold changes in the ATP demand[24,25]. The mechanisms allowing ATP homeostasis to maintain stable ATP levels is an active area of research with proposed roles for creatine phosphate, adenylate kinase, nucleotide diphosphate kinase, and calcium. Our results suggest that allosteric regulation of glycolysis is fully capable of maintaining near-constant ATP levels by itself. Another puzzling observation has been that the stable ATP maintenance mechanism is broken in pancreatic beta cells that use the sensitivity of ATP/ADP ratio to glucose levels as a readout to couple insulin secretion to blood glucose[36]. The molecular mechanism that allows for the coupling of ATP/ADP ratio to glucose levels only in beta cells has been unclear. Our results suggest that the beta-cell specific isoform of HK called Glucokinase or HK4 is responsible for this effect as it lacks feedback inhibition by G6P, which our results suggest is sufficient to break the stable ATP maintenance of glycolysis (Figure 4A).

Glycolytic enzymes are expressed at unusual levels, with GLUT, HK1, PFKP, and MCT being 10-100x less active than most of the other enzymes (Figure 6A, B). Such an expression pattern is puzzling given that glycolysis cannot proceed faster than its slowest enzymes, so 10-100x higher expression of most enzymes is seemingly wasteful. Our results provide a mechanistic explanation for this unusual expression pattern by showing that it is required to maintain high and stable ATP levels (Figure 6C).

Several organisms have modified versions of HK1 and PFKP reactions with a phosphate donor other than ATP. For example, many bacteria use PEP to phosphorylate glucose into G6P using the phosphotransferase system[37], several organisms contain ADP-dependent HK and PFK[38], and select members of all three domains of life use pyrophosphate (PPi) to phosphorylate F6P using PPi-dependent PFK[39,40]. Our results suggest that the function of these reactions might be to bypass the requirement for allosteric regulation by making the futile cycle of phosphate impossible (Figure 5A). The use of PEP or ADP makes the futile cycle impossible as PEP and ADP do not participate in the futile cycle, while the use of PPi-dependent PFK makes the futile cycle less thermodynamically favorable due to 1000-fold lower equilibrium constant of PPi-dependent PFKs vs. ATP-dependent PFKs reactions (Figure 5D, Figure 7G). Demonstration that PPi-dependent PFKs are not allosterically regulated[40] provides additional evidence that using PPi instead of ATP obviates the need for allosteric regulation.

Beyond the applications described above, our model of glycolysis can be used for a variety of applications. The model can be used to make predictions of the effect of a specific enzyme mutant on glycolysis simply by changing the kinetic parameters of the wild-type enzyme to that of the mutant enzyme. Global sensitivity analysis (Figure S3) can be used to identify kinetic parameters that control any output of glycolysis, such as the tasks of ATP homeostasis or a concentration of a specific metabolite. The rate equations for glycolytic enzymes can be used in reaction-diffusion equations to investigate the effect of the 3D distribution of glycolytic enzymes. The existence of novel regulators not included in the model can be inferred by investigating the quantitative differences between model prediction and data.

To end, we highlight that our framework of combining theory and experiments can be applied to study the regulation of any metabolic pathway. Enzyme rate equations for other metabolic pathways can be derived from *in vitro* kinetics data using the methods that we described here. Our approaches for simulations, global sensitivity analysis, and error propagation of systems of ODEs can be used for other pathways without modification. We anticipate that application of our framework to other metabolic pathways will lead to a major reevaluation of the functions of mass action and allostery in the regulation of metabolic homeostasis.

SUPPLEMENTAL INFORMATION

Supplemental information file is available for this article that includes Figures S1-S4 and Notes S1-S2. In addition, two Excel Spreadsheets with data are provided "*Data S1. Levels of enzymes, metabolites and isotope tracing.xlsx"* and *"Data S2. Kinetic rate data for glycolytic enzymes.xlsx"*.

ACKNOWLEDGEMENTS

We thank the students and instructors of the Marine Biological Laboratory course on Physical Biology of the Cell for their comments on the early version of this project; participants of the Kavli Institute for Theoretical Physics workshop on Cellular Energetics for fruitful discussions; Bradley Webb for discussions about PFK regulation; and members of Titov and Phillips labs for helpful suggestions. This research used the Savio computational cluster resource provided by the Berkeley Research Computing program at the University of California, Berkeley (supported by the UC Berkeley Chancellor, Vice Chancellor for Research, and Chief Information Officer). This work was supported in part by NIH grants DP2 GM132933 to D.V.T and 5R35 GM118043-7 to R.P. T.E. is supported by Damon Runyon Cancer Research Foundation Fellowship DRQ 01-20.

AUTHOR CONTRIBUTIONS

Conceptualization: RP, DVT

Data curation: MC, TE, DVT

Methodology: MC, TE, DVT

Investigation: MC, TE, DVT

Funding acquisition: RP, DVT

Supervision: RP, DVT

Writing – original draft: DVT

Writing – review & editing: MC, TE, RP, DVT

DECLARATION OF INTERESTS

The authors declare no competing interests.

STAR Methods

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Experimental models: Cell lines		
HeLa	ATCC	
293T	ATCC	
C2C12	ATCC	
NIH3T3	ATCC	
Software and algorithms		
Julia Programming Language	https://julialang.org	N/A
DifferentialEquations.jl package for simulation of differential equations	https://github.com/SciML/DifferentialEquations.jl	N/A
GlobalSensitivity.jl package for global sensitivity analysis	https://github.com/SciML/GlobalSensitivity.jl	N/A

RESOURCE AVAILABILITY

Lead contact
Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Denis V. Titov (titov@berkeley.edu).

Materials availability
This study did not generate new unique reagents.

Data and code availability
- All data reported in this paper is provided as supplemental excel files.
- All original code is publicly available at https://github.com/DenisTitovLab/Glycolysis.jl
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

HeLa, 293T, NIH3T3, and C2C12 cells were cultured in DMEM without pyruvate (US Biological, Cat# D9802-25L), supplemented with 25 mM glucose, 3.7 g/L NaHCO3 and 10% FBS (Gibco, Cat# 10437028).

METHOD DETAILS

Estimation of intracellular concentrations of glycolytic enzymes

We estimated the abundance of glycolytic enzymes in mammalian cell lines using publicly available proteomics data[41]. We first used proteomics data to calculate the fraction a specific glycolytic enzyme occupies compared to total proteome size as a ratio of peptide intensities for the specific glycolytic enzyme to total peptide intensity for all proteins. This approach has been shown to produce good estimates of absolute protein amounts[42]. Values of mg protein per mg total protein were then converted to concentration using the molecular weight of each enzyme and the cellular protein density of ~200 mg/mL[43]. The resulting dataset of glycolytic enzyme concentration is in *Data S1. Intracellular concentration of glycolytic enzymes, metabolites and isotope tracing data.xlsx*.

Measurement of intracellular glycolytic intermediate concentrations using LC-MS

LC/MS was used to profile and quantify the polar metabolite contents of whole-cell samples. The metabolite extraction buffer was composed of 40% methanol (Fisher Scientific, Cat# A456-4) and 40% acetonitrile (Fisher Scientific, Cat# A9554) in water (Fisher Scientific, Cat# W64) containing 0.1 M formic acid (Fisher Scientific, Cat# A11750), supplemented with a mixture of 8 isotope-labeled chemicals at 1 μM (ATP (Sigma, Cat# 710695), AMP (Sigma, Cat# 900382), glucose-6-phosphate (Cambridge Isotope Laboratories, Inc., Cat# CLM-8367) and fructose-6-phosphate (Cambridge Isotope Laboratories, Inc., Cat# CLM-8616)), 10 μM (lactate (Sigma, Cat# 490040), pyruvate (Sigma, Cat# 490709) and fructose-1,6-bisphosphate (Cambridge Isotope Laboratories, Inc., Cat# CLM-3398)) or 20 μM (glucose (Sigma, Cat# 389374)), which were used as internal standards.

Metabolite analysis was performed on an Agilent 6430 Triple Quad interfaced with an Agilent 1260 Infinity II LC system. 5 μL of each metabolite sample was injected onto a SeQuant ZIC-pHILIC column 5 μm polymer 150 X 2.1 mm (Millipore/Sigma, Cat# 1504600001). Buffer A was 100% acetonitrile; buffer B was 20 mM ammonium carbonate, 0.1% ammonium hydroxide. The chromatographic gradient was run at a flow rate of 0.150 mL/min as follows: 0-0.5 min: hold at 80% A; 0.5-30.5 min: linear gradient from 80% to 20% A; 30.5-30.8 min: hold at 20% A; 30.8-31 min: linear gradient from 20% to 80% A; 31-60 min: hold at 80% A.

The MS settings were kept consistent regardless of the chromatographic separation being tested. MS parameters were as follows: gas, 350 ºC at 11 L/min; Nebulizer, 35 psi at 4000 V. The MS was operated in negative ionization mode for all samples analyzed.

All experiments were performed in replicates of five (n=5) per sample group. Metabolite identification and quantification were performed with the Agilent MassHunter Qualitative Analysis software (version B.06.00). To confirm metabolite identities and to enable quantification, the pools of metabolite standards were used. To accurately quantify target metabolites, the final concentrations of standards were 100 μM, 50 μM, 25 μM, 12.5 μM, 6.25 μM, 3.125 μM, 1.56 μM, 781 nM, 390 nM, 195 nM, 97 nM, and 49 nM. In each sample, the raw peak area of each metabolite was divided by the raw peak area of the relevant isotope-labeled internal standard to calculate the absolute concentration. ATP, AMP, glucose-6-phosphate, fructose-1,6-bisphosphate, pyruvate, and lactate were normalized with their isotope-labeled counterparts. For determining the absolute concentrations of all other metabolites, the peak areas were normalized with 2 isotope-labeled internal standards which have the closest retention time. The raw peak area values were fit to a linear fitting curve equation, typically with $r^2 > 0.99$, which was then used to calculate the concentration of the metabolite in each metabolite extract. The final intracellular concentrations of target metabolites were then calculated from the sample dilution-fold and the corresponding cell volume which was estimated using Coulter Counter.

Two sets of isomers, fructose-6-phosphate/glucose-1-phosphate and 3-phosphoglycerate/2-phosphoglycerate were not separated under our condition and different calculation method was used to determine intracellular concentrations. For example, raw peak area of fructose-6-phosphate/glucose-1-phosphate is a sum of raw peak area of fructose-6-phosphate and glucose-1-phosphate. I determined linear fitting curve equations for each target based on standard chemicals. Since our methods targets two different transition states for most of the metabolites (in case of fructose-6-phosphate/glucose-1-phosphate, product ion value of 78.96 or 96.97), two different equations can be generated for one isomer peak. Absolute concentrations of two isomers in metabolite extract were calculated based on these two equations. Then as already explained above, the final intracellular concentrations of isomers were calculated from the sample dilution-fold and the corresponding sample cell volume.

All Whole-cell metabolite profiling was performed on ice. 1 million C2C12 or HeLa cells were seeded on 6-well plates in 2 mL of DMEM without pyruvate (US Biological, Cat# D9802-25L), supplemented with 25 mM glucose, 3.7 g/L $NaHCO_3$ and 10% FBS (Gibco, Cat# 10437028). 24 hr later, the medium was exchanged to 2 mL of DMEM medium without pyruvate, supplemented with 25 mM glucose, 3.7 g/L $NaHCO_3$, 10% FBS, and 1 μM oligomycin (MP biomedicals, Cat# 151786). After 2 hr incubation, the cell culture plate was placed on ice and the medium was aspirated. 150 μL of ice-chilled extraction buffer (40:40:20 = methanol:acetonitrile:water) containing 0.1 M formic acid and 8 isotope-labeled internal standards was added and cells were scraped with a cell lifter for 10 s. The lysate was transferred to a 1.5-mL tube on ice. 5 min later, cells were centrifuged at 17000 *g* at 4 ^{o}C for 10 min. Supernatants were collected and neutralized by adding ammonium bicarbonate (final concentration of 100 mM). These samples were analyzed with LC/MS as described above to determine intracellular concentrations of metabolites in each whole-cell extract.

For each experiment, a replicate set of cells was treated identically and used for measuring cell number and volume. The contents of each well were trypsinized and cell number and volume were measured using a Beckman Z2 Coulter Counter with a size setting of 492.8-23620 fL. Intracellular concentrations were calculated using the total number of cells per sample and the volume of each cell.

Estimates of glycolytic intermediates concentration from the literature

In addition to our own measurements of glycolytic intermediate concentrations, we have compiled a dataset of mammalian glycolytic concentrations measured under different conditions from XYZ publications. The resulting dataset combined with our own measurement is reported in Figure 2D and also provided as supplemental file *Data S1. Intracellular concentration of glycolytic enzymes, metabolites and isotope tracing data.xlsx*.

Measurement of glycolytic fluxes using [^{13}C]Glucose and [^{13}C]Lactate

C2C12 or HeLa cells (1 million) were seeded on 6-well plates in 2 mL of DMEM without pyruvate, supplemented with 25 mM glucose, 3.7 g/L $NaHCO_3$ and 10% FBS. 24 hr later, the medium was exchanged to 2 mL of DMEM medium without pyruvate, supplemented with 25 mM glucose, 3.7 g/L $NaHCO_3$, 10% FBS, and 1 μM oligomycin. After 2 hr incubation, the medium was exchanged to 2 mL of DMEM medium without pyruvate, supplemented with 25 mM $^{13}C_6$-glucose(Sigma, Cat# 389274-1G), 3.7 g/L $NaHCO_3$, 10% dialyzed FBS (ThermoFisher, Cat# 26400044), and 1 μM oligomycin or 2 mL of DMEM medium without pyruvate, supplemented with 25 mM glucose, 5 mM $^{13}C_3$-lactate (Cambridge isotope laboratory, Cat# CLM-1579-N-0.1MG), 3.7 g/L $NaHCO_3$, 10% dialyzed FBS and 1 μM oligomycin. At time points of 1, 3, 10, 30 min, cell culture plates were place on ice and medium was aspirated. Cells

were rinsed with 2 mL of ice-cold 100 mM ammonium bicarbonate for no more than 10 s and metabolites were extracted as described above using extraction buffer (40:40:20 = methanol:acetonitrile:water) containing 0.1 M formic acid and 1 μM isotopic AMP. In each sample, the raw peak area of each metabolite was divided by the raw peak area of isotope-labeled AMP internal standard to calculate the relative abundance.

Numerical simulations of differential equations constituting the Glycolysis Model

The glycolysis model is a system of ordinary differential equations (ODEs). The number of equations is equal to the number of metabolites, and each equation consists of a derivative of metabolite concentration with respect to time being equal to the sum of enzyme rates producing the metabolite minus the sum of enzyme rates consuming the metabolite:

$$\frac{d[Glucose_{media}]}{dt} = 0$$

$$\frac{d[Glucose]}{dt} = V^{GLUT} - V^{HK1}$$

$$\frac{d[G6P]}{dt} = V^{HK1} - V^{GPI}$$

$$\frac{d[F6P]}{dt} = V^{GPI} - V^{PFKP}$$

$$\frac{d[F16P]}{dt} = V^{PFKP} - V^{ALDO}$$

$$\frac{d[GAP]}{dt} = V^{ALDO} + V^{TPI} - V^{GAPDH}$$

$$\frac{d[DHAP]}{dt} = V^{ALDO} - V^{TPI}$$

$$\frac{d[BPG]}{dt} = V^{GAPDH} - V^{PGK}$$

$$\frac{d[3PG]}{dt} = V^{PGK} - V^{PGAM}$$

$$\frac{d[2PG]}{dt} = V^{PGAM} - V^{ENO} \qquad (5)$$

$$\frac{d[PEP]}{dt} = V^{ENO} - V^{PKM2}$$

$$\frac{d[Pyruvate]}{dt} = V^{PKM2} - V^{LDH}$$

$$\frac{d[Lactate]}{dt} = V^{LDH} - V^{MCT}$$

$$\frac{d[Lactate_{media}]}{dt} = 0$$

$$\frac{d[ATP]}{dt} = V^{PGK} + V^{PKM2} + V^{AK} - V^{HK1} - V^{PFKP} - V^{ATPase}$$

$$\frac{d[ADP]}{dt} = V^{HK1} + V^{PFKP} + V^{ATPase} - V^{PGK} - V^{PKM2} - V^{AK}$$

$$\frac{d[AMP]}{dt} = V^{AK}$$

$$\frac{d[Phosphate]}{dt} = V^{ATPase} - V^{GAPDH}$$

$$\frac{d[NADH]}{dt} = V^{GAPDH} - V^{LDH}$$

$$\frac{d[NAD]}{dt} = V^{LDH} - V^{GAPDH}$$

$$\frac{d[F26BP]}{dt} = 0$$

$$\frac{d[Citrate]}{dt} = 0$$

We numerically simulated the system of differential equations (5) using DifferentialEquations.jl library[28]. We typically used Rodas4 solver for stiff differential equations, absolute tolerance 10^{-12} and relative tolerance 10^{-5}. Numerical simulation of ODEs requires an initial condition consisting of metabolite concentration at time 0. We used estimates of intracellular concentrations of glycolytic intermediates as initial conditions or a steady state solution for simulations with step responses to changes in ATPase rate. A large range of initial conditions, including all metabolites being zero except for ATP, NAD, and extracellular glucose, produced similar steady-state solutions as described. We simulated differential equations for 10^{10} seconds (~3 years) to find the steady-state concentrations of metabolites. Typically, the solution was close to steady after only several seconds in accordance with a rapid rate of glycolysis observed in live cells. DifferentialEquations.jl has callback capabilities, which allowed us to change any parameter of the model (e.g., ATPase rate or concentration of any metabolite) at any time during the simulation. A typical simulation required 2-10 milliseconds to complete on a single core of Apple M1 Max processor, which allowed us to test many different conditions in a short period of time using a personal computer.

Calculation of ATP turnover, Energy released by ATP hydrolysis, Bound and Free Metabolite Concentrations

The output of the simulation of the system of differential equations (5) is the free concentrations of metabolites at every timepoint starting from initial conditions. This output can be used to calculate many properties of the glycolysis pathway such as rates of any reaction or concentrations of bound metabolites by plugging concentration into enzyme rate equations described in Note S2.

Consumption and production of ATP was calculated using the following equations:

$$ATP\ consumption = V^{ATPase} \tag{6}$$

$$ATP\ production = V^{PGK} + V^{PKM2} + V^{AK} - V^{HK1} - V^{PFKP} \tag{7}$$

Energy released by ATP hydrolysis in units of k_BT was calculated using the following equation:

$$E^{ATPase} = -ln\left(\frac{\Gamma}{K_{eq}}\right) = -ln\left(\frac{\frac{[ADP][Phosphate]}{[ATP]}}{K_{eq}}\right) \quad (8)$$

where Γ is mass action ratio of ATP hydrolysis and K_{eq} is equilibrium constant of ATP hydrolysis.

The concentration of enzyme-bound metabolites was calculated using binding equations described in the Supplementary Text section. Accounting for enzyme-bound metabolites is essential to ensure an accurate comparison of model prediction to whole-cell measurements as the latter represents a sum of enzyme-bound and free metabolites. Intracellular concentrations as glycolytic enzymes are 1-10 μM, which is comparable to intracellular concentrations of some glycolytic intermediates. Therefore, some of the metabolites are mostly present in an enzyme-bound state, and their concentrations will be underestimated by the model without accounting for the enzyme-bound fraction.

Estimation of Model Uncertainty

Outputs of the glycolysis model have uncertainty due to uncertainty of estimates of model parameters such as kinetic constants, thermodynamic constants, and enzyme concentrations. Understanding the uncertainty of model predictions is important for quantitative comparisons of model predictions to data and for understanding the robustness of model predictions. We used bootstrapping to calculate the uncertainty of glycolysis model predictions. We rerun the model 10,000 times using different sets of parameter values, where each parameter value is generated by randomly drawing a value from a normal distribution with a mean and standard deviation corresponding to the mean and standard deviation of that parameter estimate. We bootstrapped all the parameters except we kept the kinetic parameters of ATPase and Adenylate Kinase unchanged during bootstrapping as these enzymes are not part of glycolysis and set Vmax of ATPase to be equal to a fraction of Vmax of HK1 or Vmax PFKP whichever was smaller in a particular bootstrapped parameter set. In most analyses, we only reported the data from bootstrapped simulations that could match ATP supply and demand as indicated in relevant figure legends because including simulations that come to equilibrium due to inability to produce ATP fast enough is not informative as they represent a completely different state of the glycolysis system akin to a dead cell. In summary, this general bootstrapping approach allows us to estimate the uncertainty of any output of our model.

Global Sensitivity Analysis

Sensitivity analysis is a set of powerful approaches aimed to systematically explore the role of specific parameters of an ODE model in controlling specific outputs of the model, such as maintenance of high ATP level at a range of ATP turnovers, prediction of metabolite concentration, speed of the response to perturbation, etc. Sensitivity analysis can be broadly subdivided into local sensitivity analysis and global sensitivity analysis methods. Local sensitivity analysis methods using partial derivates were historically more popular especially for simple models where analytical derivatives can be calculated. When applied to models of metabolic pathways, local sensitivity analysis is sometimes referred to as metabolic control theory. The limitations of local sensitivity analysis are that it only considers the effect of one parameter at a time (i.e., ignores interactions between parameters) and only estimates the role of the parameter at a specific value of that parameter and every other parameter (i.e., local). To overcome these limitations, global sensitivity analysis methods were developed that allow systematic evaluation of parameter importance at a large range of parameter values (i.e., global)

and that can estimate the effect of interactions between parameters. The only limitation of variance decomposition methods is that they are much more computationally intensive than local sensitivity analysis methods, which limited their use before early 2000s when the computational power have caught up to the demands of these methods.

We used a global sensitivity analysis method called variance decomposition to learn about the role of specific model parameters in controlling specific outputs of our model. Here we briefly describe the idea behind this method and refer the reader to several excellent books that have been written about global sensitivity analysis for more in-depth discussion[34,44]. First, we chose a scalar output *Y* of the model that we want to investigate. We focused on investigating the model parameters required to maintain high ATP concentration, high energy of ATP hydrolysis, and the ability of model to match ATP production and ATP consumption rates. Then we perform repeated simulations of the model using different starting parameters drawn from a uniform distribution with values 3-fold lower to 3-fold higher than the default model parameters. We then calculated how much does each parameter or combination of parameters contribute to the variance of output *Y* observed during repeated simulations. A parameter that is important for controlling *Y* will have large contribution and the parameter that is not important will have small or no contributions. Variance of output *Y* that depends on k input parameters *P* could be uniquely decomposed into the following terms:

$$Var(Y) = \sum_i V_i + \sum_i \sum_{j>i} V_{ij} + \sum_i \sum_{j>i} \sum_{l>j} V_{ijl} + \cdots + V_{1,2\ldots k} \tag{9}$$

where

$$V_i = Var(E(Y|P_i))$$

$$V_{i,j} = Var\left(E\left(Y\middle|P_i, P_j\right)\right) - V_i - V_j$$

$$V_{i,j,l} = Var\left(E\left(Y\middle|P_i, P_j, P_l\right)\right) - V_i - V_j - V_l - V_{i,j} - V_{i,l} - V_{j,l}$$

Two types of sensitivity indexes, S_i and S_i^T, are useful in analyzing the contribution of input parameters. S_i is called sensitivity index, importance measure or first-order effect:

$$S_i = \frac{V_i}{Var(Y)} \tag{10}$$

Another way to understand S_i is that it is a fraction of the variance of *Y* that would disappear if we fixed parameter P_i at some value. Parameter P_i would be considered important for controlling output Y if fixing its value would lead to a large reduction of the variance of Y (i.e., S_i is large). S_i has several nice properties where for an additive model $\sum S_i = 1$ and for any model with independent input parameters $\sum S_i \leq 1$. For example, a large value of $1 - \sum S_i$ indicates that interactions between parameters are important for controlling a specific output of the model. More generally, for any model with independent input parameters:

$$\sum_i S_i + \sum_i \sum_{j>i} S_{ij} + \sum_i \sum_{j>i} \sum_{l>j} S_{ijl} + \cdots + S_{1,2\ldots k} = 1 \tag{11}$$

S_i^T is called the total-order effect and contains all terms in equation (11) that contain *i*:

$$S_i^T = 1 - \frac{Var(E(Y|P_{-i}))}{Var(Y)} = \frac{E(Var(Y|P_{-i}))}{Var(Y)} = S_i + \sum_{j\neq i} S_{ij} + \sum_{l\neq i,j} \sum_{j\neq i} S_{ijl} + \cdots \tag{12}$$

Another way to understand S_i^T is that it is a fraction of the variance of Y that would be left if we fixed all parameters except parameter P_i at some value. Parameter P_i would be considered important for controlling output Y if fixing all the parameters except P_i at some value would still leave a large fraction of variance of Y (i.e., S_i^T is large). Unlike S_i, $\sum S_i^T$ can be larger or smaller than 1 and equation (11) does not hold for S_i^T because S_i^T and S_j^T would have one or several overlapping terms (e.g., S_{ij}).

We have calculated S_i and S_i^T for various outputs of the model using the Sobol method of the GlobalSensitivity.jl package.

Adoption of the Glycolysis Model for Simulation of [U-^{13}C]-Glucose and [U-^{13}C]-Lactate Tracing

Glycolysis model can be used to predict the results of heavy isotope tracing experiment. To predict tracing patterns, we decompose each rate equation into enzyme activity part and mass action part as described previously[35,45,46]:

$$V = Enzyme\ Activity \cdot \left([S_1]\cdot[S_2] - \frac{[P_1]\cdot[P_2]}{Keq}\right) \tag{13}$$

Enzyme activity depends on the total concentration of metabolites (i.e., sum of heavy and light isomers) while the mass action part is responsible for the propagation of isotope label. Therefore, for modeling of isotope tracing we modify all kinetic equation to be able to heavy and light isotope labeled metabolites that we incorporate in mass action part of all the equations, and we double the number of ODEs to account for time dependent changes of both heavy and light isomers (i.e., we have one ODE for each isotope).

Kinetic equations and numerical simulation of differential equations constituting the Two-Enzyme Model

We described the kinetics of Enzyme 1, Enzyme 2, and ATPase using simplified rate equations. We started from reversible Michaelis-Menten equations[35] for Enzyme 2 and ATPase. For Enzyme 1, we used the MWC model to add allosteric activation by ADP and phosphate and allosteric inhibition by ATP. We assumed that ADP and Phosphate only bind to active conformation, ATP only binds to inactive conformation, inactive MWC conformation of Enzyme 1 is catalytically inactive. The resulting rate equations for Enzyme 1, Enzyme 2 and ATPase are:

$$V^{Enz\,1} = \frac{V_{max}^{Enz\,1}}{\left(1 + \frac{L\left(1+\frac{[ATP]}{K_A^{ATP}}\right)^n}{\left(1+\frac{[ADP]}{K_A^{ADP}}\right)^n\left(1+\frac{[Phosphate]}{K_A^{Phosphate}}\right)^n}\right)}\left(\frac{\frac{[G][ATP]}{K_M^G K_M^{ATP}}}{\left(1+\frac{[G][ATP]}{K_M^G K_M^{ATP}}+\frac{[X][ADP]}{K_M^X K_M^{ADP}}\right)}\right)\left(1-\frac{1}{K_{eq}}\frac{[X][ADP]}{[G][ATP]}\right) \tag{14}$$

$$V^{Enz\,2} = V_{max}^{Enz\,2}\left(\frac{\frac{[X][ADP]^2[Phosphate]}{K_M^X(K_M^{ADP})^2K_M^{Phosphate}}}{1+\frac{[X][ADP]^2[Phosphate]}{K_M^X(K_M^{ADP})^2K_M^{Phosphate}}+\frac{[L][ATP]^2}{K_M^L(K_M^{ATP})^2}}\right)\left(1-\frac{1}{K_{eq}}\frac{[L][ATP]^2}{[X][ADP]^2[Phosphate]}\right) \tag{15}$$

$$V^{ATPase} = V_{max}^{ATPase}\left(\frac{\frac{[ATP]}{K_M^{ATP}}}{\left(1+\frac{[ATP]}{K_M^{ATP}}+\frac{[Phosphate][ADP]}{K_M^{Phosphate}K_M^{ADP}}\right)}\right)\left(1-\frac{1}{K_{eq}^{ATPase}}\frac{[ADP][Phosphate]}{[ATP]}\right) \tag{16}$$

We then assumed that catalytic sites are saturated with substrates and regulatory sites are saturated with regulators simplifying the rate equations to:

$$V^{Enz\,1} = \frac{\frac{V_{max}^{Enz\,1}}{K_M^G K_M^{ATP}}}{\left(1 + \left(\frac{L[ATP]}{[ADP][Phosphate]}\right)^n\right)} \left([G][ATP] - \frac{1}{K_{eq}}[X][ADP]\right) \tag{17}$$

$$V^{Enz\,2} = \frac{V_{max}^{Enz\,2}}{K_M^X (K_M^{ADP})^2 K_M^{Phosphate}} \left([X][ADP]^2[Phosphate] - \frac{1}{K_{eq}}[L][ATP]^2\right) \tag{18}$$

$$V^{ATPase} = V_{max}^{ATPase}\left(1 - \frac{1}{K_{eq}^{ATPase}}\frac{[ADP][Phosphate]}{[ATP]}\right) \tag{19}$$

The two-enzyme model is a system of ordinary differential equations (ODEs). The number of equations is equal to the number of metabolites, and each equation consists of a derivative of metabolite concentration with respect to time being equal to the sum of enzyme rates producing the metabolite minus the sum of enzyme rates consuming the metabolite:

$$\begin{aligned}
\frac{d[G]}{dt} &= 0 \\
\frac{d[X]}{dt} &= V^{Enzyme\,1} - V^{Enzyme\,2} \\
\frac{d[L]}{dt} &= 0 \\
\frac{d[ATP]}{dt} &= -V^{Enzyme\,1} + 2V^{Enzyme\,2} - V^{ATPase} \\
\frac{d[ADP]}{dt} &= V^{Enzyme\,1} - 2V^{Enzyme\,2} + V^{ATPase} \\
\frac{d[Phosphate]}{dt} &= V^{ATPase} - V^{Enzyme\,2}
\end{aligned} \tag{20}$$

We numerically simulated the system of differential equations (5) using DifferentialEquations.jl library[28] using the same approaches as for the full model described above.

REFERENCES

1. DeBerardinis, R.J., and Chandel, N.S. (2020). We need to talk about the Warburg effect. Nature Metabolism *2*, 127–129. 10.1038/s42255-020-0172-2.

2. King, M.P., and Attardi, G. (1989). Human cells lacking mtDNA: repopulation with exogenous mitochondria by complementation. Science *246*, 500–503. 10.1126/science.2814477.

3. Titov, D.V., Cracan, V., Goodman, R.P., Peng, J., Grabarek, Z., and Mootha, V.K. (2016). Complementation of mitochondrial electron transport chain by manipulation of the NAD+/NADH ratio. Science *352*, 231–235. 10.1126/science.aad4017.

4. Sel'kov, E.E. (1975). Stabilization of energy charge, generation of oscillations and multiple steady states in energy metabolism as a result of purely stoichiometric regulation. Eur J Biochem *59*, 151–157. 10.1111/j.1432-1033.1975.tb02436.x.

5. Rapoport, T.A., Otto, M., and Heinrich, R. (1977). An extended model of the glycolysis in erythrocytes. Acta Biol Med Ger *36*, 461–468.

6. Termonia, Y., and Ross, J. (1981). Oscillations and control features in glycolysis: numerical analysis of a comprehensive model. Proc Natl Acad Sci U S A *78*, 2952–2956. 10.1073/pnas.78.5.2952.

7. Joshi, A., and Palsson, B.O. (1990). Metabolic dynamics in the human red cell. Part IV--Data prediction and some model computations. J Theor Biol *142*, 69–85. 10.1016/s0022-5193(05)80013-x.

8. Teusink, B., Bakker, B.M., and Westerhoff, H.V. (1996). Control of frequency and amplitudes is shared by all enzymes in three models for yeast glycolytic oscillations. Biochim Biophys Acta *1275*, 204–212. 10.1016/0005-2728(96)00026-6.

9. Mulquiney, P.J., and Kuchel, P.W. (1999). Model of 2,3-bisphosphoglycerate metabolism in the human erythrocyte based on detailed enzyme kinetic equations: equations and parameter refinement. Biochem J *342 Pt 3*, 581–596.

10. Teusink, B., Passarge, J., Reijenga, C.A., Esgalhado, E., van der Weijden, C.C., Schepper, M., Walsh, M.C., Bakker, B.M., van Dam, K., Westerhoff, H.V., et al. (2000). Can yeast glycolysis be understood in terms of in vitro kinetics of the constituent enzymes? Testing biochemistry. Eur J Biochem *267*, 5313–5329. 10.1046/j.1432-1327.2000.01527.x.

11. Hynne, F., Danø, S., and Sørensen, P.G. (2001). Full-scale model of glycolysis in Saccharomyces cerevisiae. Biophys Chem *94*, 121–163. 10.1016/s0301-4622(01)00229-0.

12. Chandra, F.A., Buzi, G., and Doyle, J.C. (2011). Glycolytic oscillations and limits on robust efficiency. Science *333*, 187–192. 10.1126/science.1200705.

13. Eunen, K. van, Kiewiet, J.A.L., Westerhoff, H.V., and Bakker, B.M. (2012). Testing Biochemistry Revisited: How In Vivo Metabolism Can Be Understood from In Vitro

Enzyme Kinetics. PLOS Computational Biology *8*, e1002483. 10.1371/journal.pcbi.1002483.

14. Smallbone, K., Messiha, H.L., Carroll, K.M., Winder, C.L., Malys, N., Dunn, W.B., Murabito, E., Swainston, N., Dada, J.O., Khan, F., et al. (2013). A model of yeast glycolysis based on a consistent kinetic characterisation of all its enzymes. FEBS Lett *587*, 2832–2841. 10.1016/j.febslet.2013.06.043.

15. van Heerden, J.H., Wortel, M.T., Bruggeman, F.J., Heijnen, J.J., Bollen, Y.J.M., Planqué, R., Hulshof, J., O'Toole, T.G., Wahl, S.A., and Teusink, B. (2014). Lost in transition: start-up of glycolysis yields subpopulations of nongrowing cells. Science *343*, 1245114. 10.1126/science.1245114.

16. Shestov, A.A., Liu, X., Ser, Z., Cluntun, A.A., Hung, Y.P., Huang, L., Kim, D., Le, A., Yellen, G., Albeck, J.G., et al. (2014). Quantitative determinants of aerobic glycolysis identify flux through the enzyme GAPDH as a limiting step. Elife *3*. 10.7554/eLife.03342.

17. Mulukutla, B.C., Yongky, A., Daoutidis, P., and Hu, W.-S. (2014). Bistability in Glycolysis Pathway as a Physiological Switch in Energy Metabolism. PLOS ONE *9*, e98756. 10.1371/journal.pone.0098756.

18. Monod, J., Wyman, J., and Changeux, J.-P. (1965). On the nature of allosteric transitions: A plausible model. Journal of Molecular Biology *12*, 88–118. 10.1016/S0022-2836(65)80285-6.

19. Koshland, D.E., Némethy, G., and Filmer, D. (1966). Comparison of Experimental Binding Data and Theoretical Models in Proteins Containing Subunits*. Biochemistry *5*, 365–385. 10.1021/bi00865a047.

20. Marzen, S., Garcia, H.G., and Phillips, R. (2013). Statistical Mechanics of Monod–Wyman–Changeux (MWC) Models. Journal of Molecular Biology *425*, 1433–1460. 10.1016/j.jmb.2013.03.013.

21. Einav, T., Mazutis, L., and Phillips, R. (2016). Statistical Mechanics of Allosteric Enzymes. J Phys Chem B *120*, 6021–6037. 10.1021/acs.jpcb.6b01911.

22. Phillips, R. (2020). The Molecular Switch (Princeton University Press).

23. Milo, R., and Phillips, R. (2015). Cell Biology by the Numbers 1st edition. (Garland Science).

24. Balaban, R.S., Kantor, H.L., Katz, L.A., and Briggs, R.W. (1986). Relation between work and phosphate metabolite in the in vivo paced mammalian heart. Science *232*, 1121–1123. 10.1126/science.3704638.

25. Allen, P.S., Matheson, G.O., Zhu, G., Gheorgiu, D., Dunlop, R.S., Falconer, T., Stanley, C., and Hochachka, P.W. (1997). Simultaneous 31P MRS of the soleus and gastrocnemius in Sherpas during graded calf muscle exercise. Am J Physiol *273*, R999-1007. 10.1152/ajpregu.1997.273.3.R999.

26. Barkai, N., and Leibler, S. (1997). Robustness in simple biochemical networks. Nature *387*, 913–917. 10.1038/43199.

27. Beber, M.E., Gollub, M.G., Mozaffari, D., Shebek, K.M., Flamholz, A.I., Milo, R., and Noor, E. (2022). eQuilibrator 3.0: a database solution for thermodynamic constant estimation. Nucleic Acids Res *50*, D603–D609. 10.1093/nar/gkab1106.

28. DifferentialEquations.jl – A Performant and Feature-Rich Ecosystem for Solving Differential Equations in Julia https://openresearchsoftware.metajnl.com/articles/10.5334/jors.151/.

29. Bezanson, J., Edelman, A., Karpinski, S., and Shah, V.B. (2015). Julia: A Fresh Approach to Numerical Computing. arXiv:1411.1607 [cs].

30. Transtrum, M.K., Machta, B.B., Brown, K.S., Daniels, B.C., Myers, C.R., and Sethna, J.P. (2015). Perspective: Sloppiness and emergent theories in physics, biology, and beyond. J. Chem. Phys. *143*, 010901. 10.1063/1.4923066.

31. Heinrich, L., Bennett, D., Ackerman, D., Park, W., Bogovic, J., Eckstein, N., Petruncio, A., Clements, J., Pang, S., Xu, C.S., et al. (2021). Whole-cell organelle segmentation in volume electron microscopy. Nature *599*, 141–146. 10.1038/s41586-021-03977-3.

32. Luby-Phelps, K. (2000). Cytoarchitecture and physical properties of cytoplasm: volume, viscosity, diffusion, intracellular surface area. Int Rev Cytol *192*, 189–221. 10.1016/s0074-7696(08)60527-6.

33. Ellis, R.J. (2001). Macromolecular crowding: obvious but underappreciated. Trends in Biochemical Sciences *26*, 597–604. 10.1016/S0968-0004(01)01938-7.

34. Saltelli, A. (2008). Global Sensitivity Analysis: The Primer 1st edition. (Wiley-Interscience).

35. Noor, E., Flamholz, A., Liebermeister, W., Bar-Even, A., and Milo, R. (2013). A note on the kinetics of enzyme action: a decomposition that highlights thermodynamic effects. FEBS Lett *587*, 2772–2777. 10.1016/j.febslet.2013.07.028.

36. Newgard, C.B., and McGarry, J.D. (1995). Metabolic coupling factors in pancreatic beta-cell signal transduction. Annu Rev Biochem *64*, 689–719. 10.1146/annurev.bi.64.070195.003353.

37. Deutscher, J., Francke, C., and Postma, P.W. (2006). How phosphotransferase system-related protein phosphorylation regulates carbohydrate metabolism in bacteria. Microbiol Mol Biol Rev *70*, 939–1031. 10.1128/MMBR.00024-06.

38. Guixé, V., and Merino, F. (2009). The ADP-dependent sugar kinase family: kinetic and evolutionary aspects. IUBMB Life *61*, 753–761. 10.1002/iub.217.

39. Reeves, R.E., South, D.J., Blytt, H.J., and Warren, L.G. (1974). Pyrophosphate:D-fructose 6-phosphate 1-phosphotransferase. A new enzyme with the glycolytic function of 6-phosphofructokinase. J Biol Chem *249*, 7737–7741.

40. Siebers, B., Klenk, H.P., and Hensel, R. (1998). PPi-dependent phosphofructokinase from Thermoproteus tenax, an archaeal descendant of an ancient line in phosphofructokinase evolution. J Bacteriol *180*, 2137–2143. 10.1128/JB.180.8.2137-2143.1998.

41. Geiger, T., Wehner, A., Schaab, C., Cox, J., and Mann, M. (2012). Comparative proteomic analysis of eleven common cell lines reveals ubiquitous but varying expression of most proteins. Mol Cell Proteomics *11*, M111.014050. 10.1074/mcp.M111.014050.

42. Wiśniewski, J.R., Hein, M.Y., Cox, J., and Mann, M. (2014). A "proteomic ruler" for protein copy number and concentration estimation without spike-in standards. Mol Cell Proteomics *13*, 3497–3506. 10.1074/mcp.M113.037309.

43. Milo, R. (2013). What is the total number of protein molecules per cell volume? A call to rethink some published values. Bioessays *35*, 1050–1055. 10.1002/bies.201300066.

44. Saltelli, A., Tarantola, S., Campolongo, F., and Ratto, M. (2004). Sensitivity Analysis in Practice: A Guide to Assessing Scientific Models (Halsted Press).

45. Reich, J., and Selkov, E. (1981). Reich J. G. and Selkov E. E., Energy Metabolism of the Cell. Academic Press, London, 1981.

46. Hofmeyr, J.H. (1995). Metabolic regulation: a control analytic perspective. J Bioenerg Biomembr *27*, 479–490. 10.1007/BF02110188.

SUPPLEMENTAL FIGURES

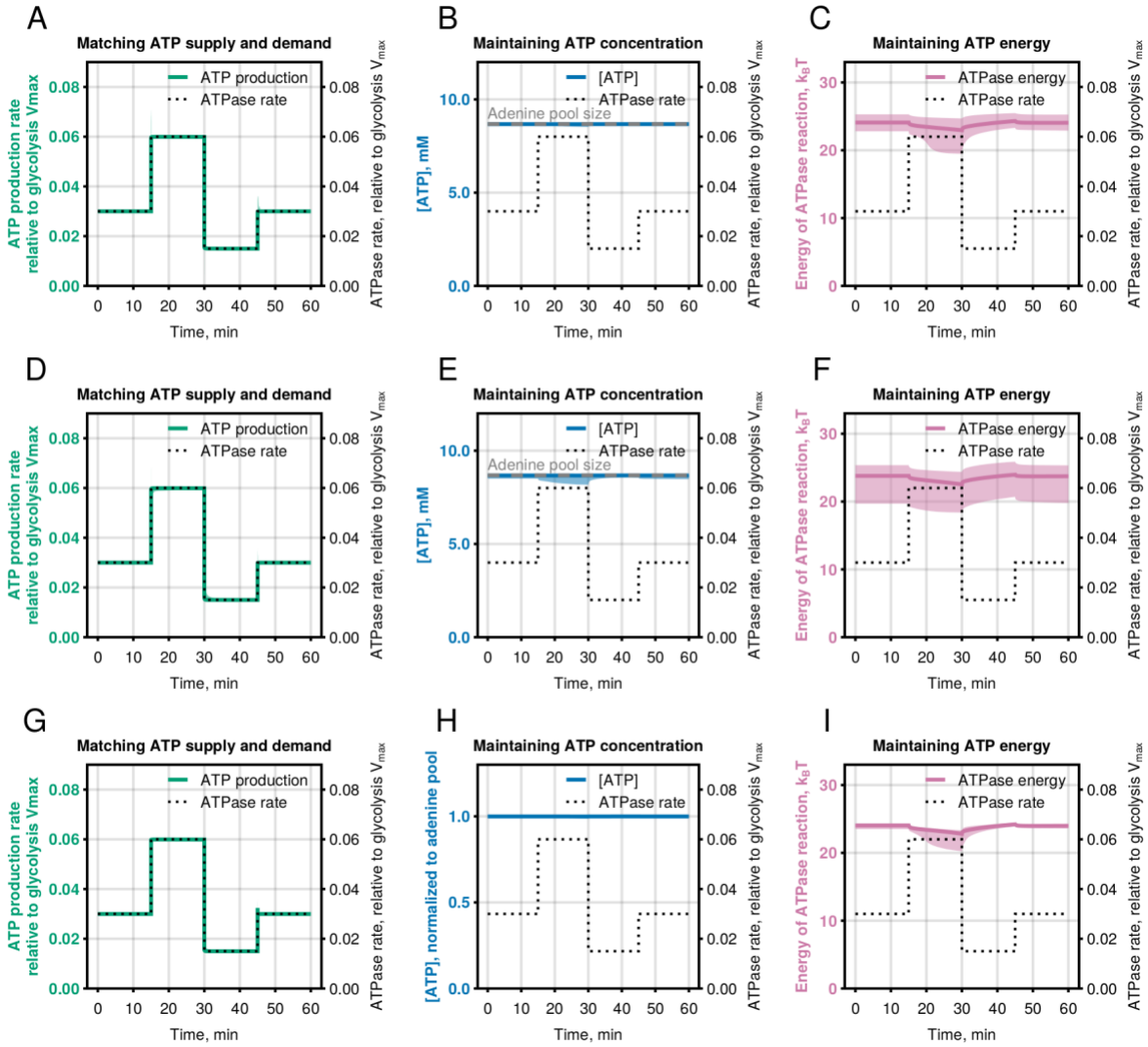

Figure S1. Related to Figure 2. Sensitivity of model predictions to enzyme levels and experimental error in biochemical constants, protein and cofactor pool concentrations, and initial concentrations of intermediates. (A-I) Model simulations showing median changes and 95% confidence interval of (A, D, G) ATP production rate, (B, E, H) ATP concentration (total adenine pool size is labeled with dashed grey line), and (C, F, I) energy released during ATP hydrolysis in response to a 2-fold stepwise increase or decrease of ATPase rate from 10,000 bootstrapped models with (A-C) enzyme concentrations drawn from a uniform distribution spanning 9-fold range around the default model value, (D-F) parameters drawn from a normal distribution with mean and SD being equal to experimental mean and SEM for each parameter of the model, (G-I) initial concentration of metabolites taken from a normal distribution with mean and SD being equal to experimental mean and SEM for each metabolite level. Blue line shows median value at each ATPase rate and blue ribbon is 95% confidence interval. ATPase energy is calculated as a natural logarithm of the disequilibrium ratio for the ATPase reaction (i.e., mass-action ratio divided by the equilibrium constant) for the ATPase reaction. Julia code to reproduce this figure is available at https://github.com/DenisTitovLab/Glycolysis.jl.

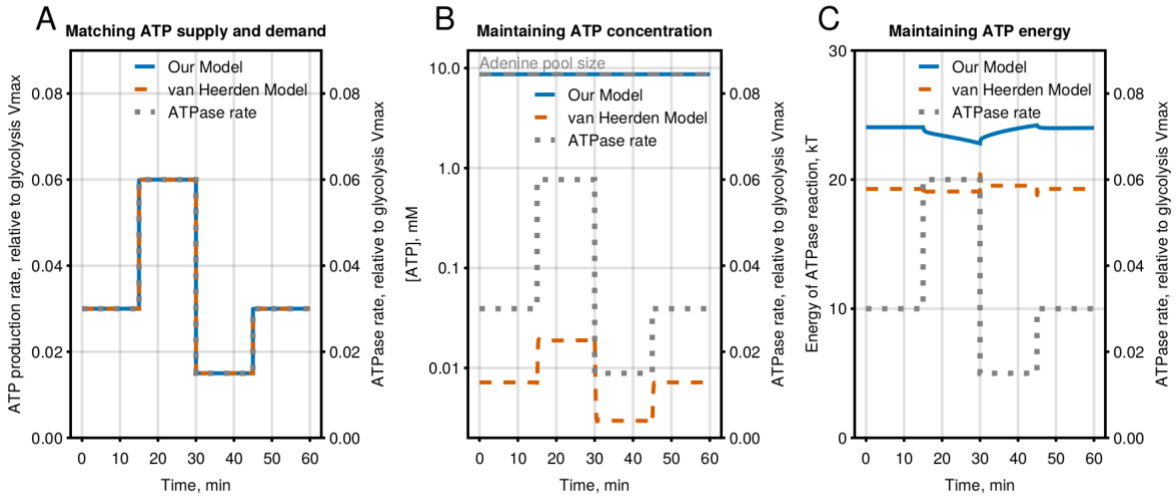

Figure S2. Related to Figure 2. Comparison of our glycolysis model with van Heerden model. (A-C) Comparison of our model and van Heerden model's ability to (A) match ATP supply and demand, (B) maintain ATP concentration, and (C) maintain energy released during ATP hydrolysis in response to a 2-fold stepwise increase or decrease of ATPase rate. Same conditions as our model in Figure 2A-C. Energy is calculated as a natural logarithm of the ratio of mass balance to the thermodynamic equilibrium constant Keq for the ATPase reaction. The total adenine pool size (i.e., ATP + ADP + AMP) is highlighted with a dashed grey line in (B). Julia code to reproduce this figure is available at https://github.com/DenisTitovLab/Glycolysis.jl.

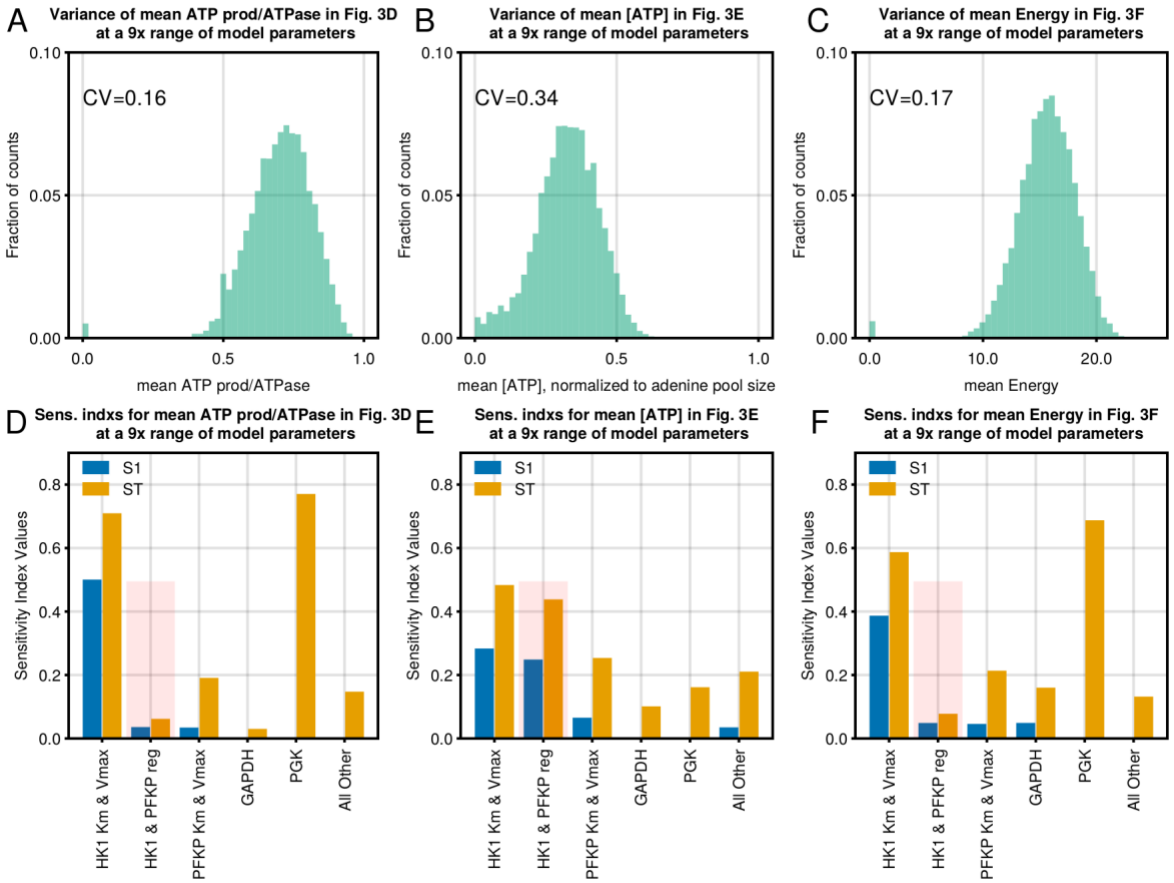

Figure S3. Related to Figure 3. Global sensitivity analysis of the model's ability to maintain ATP level, ATPase energy, and ATP production. (A-C) Distribution of (A) mean ATP production to ATPase rate ratios, (B) mean ATP concentrations normalized by adenine pool size, (C) mean ATP hydrolysis energies from a simulation like in Figure 3D-F, respectively, repeated 10,000 times using model parameters drawn from a uniform distribution spanning 9-fold range around the default model value for each parameter. **(D-F)** Sensitivity indexes of glycolysis model parameters that explain the variance of (D) mean ATP production to ATPase rate ratios, (E) mean ATP concentrations normalized by adenine pool size, (F) mean ATP hydrolysis energies, shown in Panels (A-C), respectively. Julia code to reproduce this figure is available at https://github.com/DenisTitovLab/Glycolysis.jl.

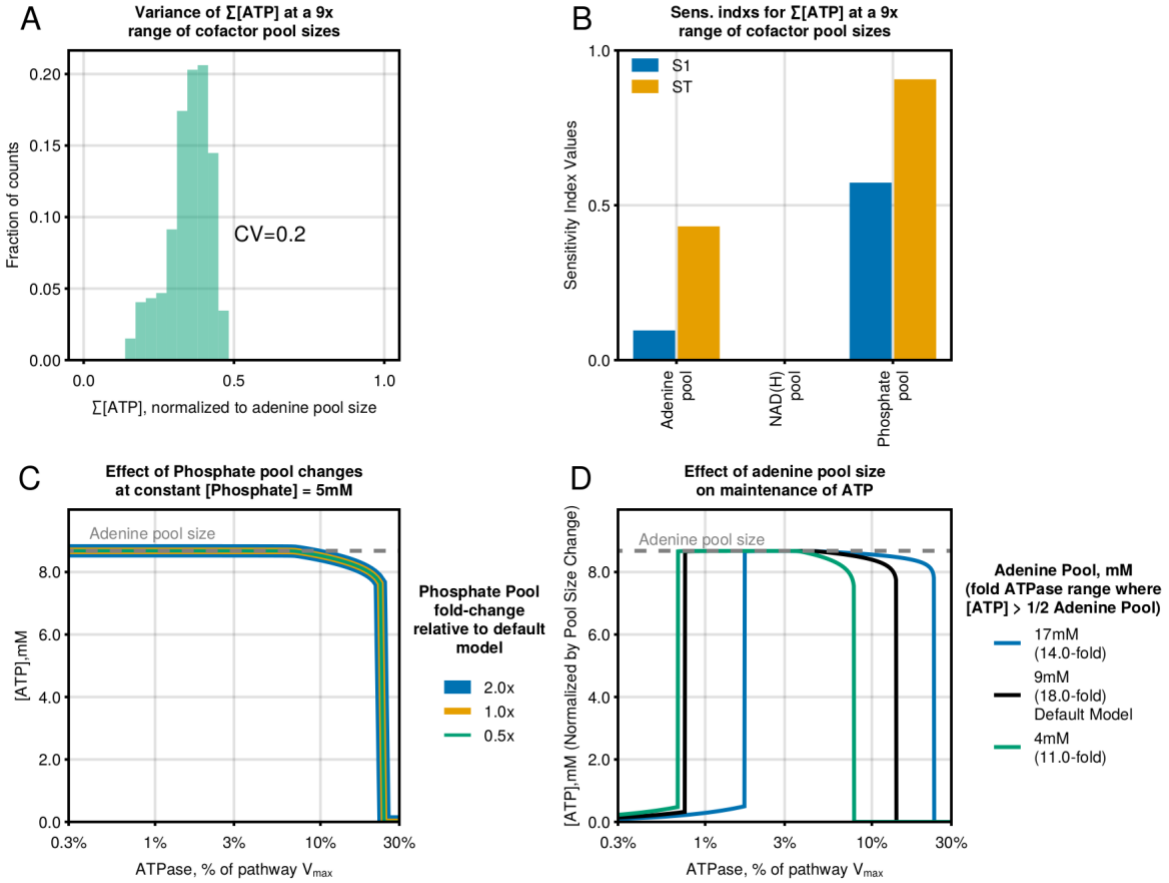

Figure S4. Related to Figure 6. Global sensitivity analysis of the effect of cofactor pool sizes on ATP level maintenance. (A) The variance of the sum of the ratio of ATP to ATP+ADP at ATPase V_{max} range of 1-50% of HK1 V_{max} from 1,000 simulations with a randomly drawn 9-fold range of Adenine (i.e., ATP + ADP + AMP), NAD(H) (i.e., NAD^+ + NADH), and Phosphate (i.e., sum of all phosphate groups except for ATP, ADP, and AMP) pool sizes around the default model values. **(B)** Sensitivity indexes of two enzyme model parameters that explain the variance of ratio of ATP to ATP+ADP shown in Panel A. **(C)** Effect of changes of Phosphate pool size on models ability to maintain ATP levels at a fixed concentration of Phosphate = 5 mM. **(D)** Steady-state ATP concentrations at a range of ATPase rates of the glycolysis model with different levels of adenine pool size. The total adenine pool size (i.e., ATP + ADP + AMP) is highlighted with a dashed grey line. Julia code to reproduce this figure is available at https://github.com/DenisTitovLab/Glycolysis.jl.